

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 13, 2002, 17:36:08 ; Search time 1833.38 Seconds

(without alignments)
11425.598 Million cell updates/sec

Title: US-09-761-580-1_COPY_800_1800

Perfect score: 1001
Sequence: 1 atatacagcattgctgac.....cagttgagccagtgctt 1001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

11 number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: gb_ba:*
2: gb_hlg:*
3: gb_la:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
------------	-------	-------------	--------	-------	-------------

1	1001	100.0	2583	6	AX151744	AX151744 Sequence
2	1001	100.0	2583	9	HSPDCE2	Y00978 Human mRNA
3	1001	100.0	2696	9	AK057299	AK057299 Homo sapi
4	961	96.0	2540	9	H0MIGMBC	J03866 Homo sapien
5	871.4	87.1	2543	4	AB036739	AB036739 Sus scrofa
6	768.6	76.8	1700	10	AY044265	AY044265 Mus muscu
7	747.8	74.7	2266	10	RATPDCE2	D10655 Rattus norv
8	735.8	73.5	1370	10	RATPDCE2	D00092 Rattus norv
9	735.8	73.5	1370	10	RATPDCE2	M16075 Rat Primary
10	735.8	73.5	1370	10	RATPDCE2	AY05169 Synthetic p
11	701	70.0	966	6	A05169	X13822 H sapiens m
12	655.8	65.5	660	9	HSDHACTYL	BC009202 Mus muscu
13	264	26.4	2038	10	BC009202	AX03511 Sequence
14	223	22.3	319	6	AX03511	U06634 Dictyostell
15	210	21.0	1833	3	DDU06634	AL023395 S. pombe c
16	205.8	20.6	40897	8	SPCC794	AY033001 Arabidops
17	202.2	20.2	1620	8	AY033001	AF367302 Arabidops
18	200.8	20.1	2001	8	AF367302	AF135014 Zea mays
19	186.8	18.7	1981	8	AF135014	Z77659 Caenorhabdi
20	178.4	17.8	2797	3	CER23B12	AC084541 Caenorhab
21	170.2	17.0	40961	3	CER23B12	AX341497 Sequence
22	169.2	16.9	172	6	AX341497	X93605 Zymomonas m
23	168.4	16.8	8238	1	ZMPDHBGEN	AP002502 Homo sapi
24	167	16.7	161920	2	AP002502	AP002007 Homo sapi
25	167	16.7	168922	2	AP002007	AP000788 Homo sapi
26	167	16.7	175419	2	AP000788	AC009648 Homo sapi
27	167	16.7	187516	2	AC009648	AE008633 Rickettsi
28	162	16.2	11199	1	AE008633	Y13145 Homo sapien
29	162	16.2	279110	1	RPRX03	U79296 Human dlynd
30	157.6	15.7	1744	9	HSPDCE2	AF001437 Homo sapi
31	157.6	15.7	1949	9	HSPDCE2	AP000788 Homo sapi
32	157.6	15.7	2320	9	AP000788	BC010389 Homo sapi
33	156.8	15.7	175419	2	AP000788	U08238 Homo sapien
34	156	15.6	2250	9	HS082328	AC025500 Mus muscu
35	156	15.6	2365	2	AC025500	AC027079 Mus muscu
36	141.2	14.1	74587	2	AC027079	AC094189 Rattus no
37	141.2	14.1	110931	2	AC094189	AE009644 Brucella
38	140.8	14.1	62280	2	AE009644	Z46230 A. thaliana
39	134.8	13.5	9353	1	ATDHPATF	AF066080 Arabidops
40	127.2	12.7	2175	8	ATDHPATF	AC017138 Drosophil
41	127.2	12.7	2274	8	AC017138	AC008354 Drosophil
42	125.8	12.6	41309	2	AC008354	AE003617 Drosophil
43	125.8	12.6	159350	3	AE003617	AL114433 Borrelia
44	125.8	12.6	257224	3	CNS01BMX	
45	124.6	12.4	660	8	CNS01BMX	

ALIGNMENTS

RESULT 1
LOCUS AX151744 2583 bp DNA linear PAT 22-JUN-2001
DEFINITION Sequence 1 from Patent EP1118679.
ACCESSION AX151744
VERSION AX151744.1 GI:14533673
KEYWORDS
SOURCE
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
1 (bases 1 to 2583)
AUTHORS Smith,J.C., Anand,R. and Morten,J.E.
TITLE Method for diagnosing polymorphisms in the human pth e2 gene
JOURNAL Patent: EP 1118679-A 1 25-JUN-2001;

FEATURES
Location/Qualifiers
source
1..2583
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 742 a 562 c 564 g 715 t
ORIGIN

```

Query Match      100.0%; Score 1001; DB 6; Length 2583;
Best Local Similarity 100.0%; Pred. No. 2e-234;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atatacagcatttgctgactaagcccaacgaaglaacagatttaaacacacaagtc 60
DB 800 ATATATACGATTGTGCTGACTATAGGCCAACCAAGTAACAGATTAAACCAAGTGC 859
QY 61 caccactaccacccccggggtggtcgtgttctcccaactcccccagcttagctccta 120
DB 860 CACCACCTACCCACCCCGGGTGGCGGTCTCTCAACTCCCGCCTTAAGTCTCTTA 919
QY 121 caccctcaacacccctgcacagctactccctgagccaaagggaagggtgtgttagcc 180
DB 920 CACCTTCAGACACCTGCCACAGTACTCCGTGTGACCAAGGAAGGAGTGTGTTAGCC 979
QY 181 ctcttgcaagaagcttgagcagtagaagaaggatgtatcttaccagaagtaaaaggagc 240
DB 980 CTCTTCAGAGAGAGTGGCAGTAGAGAAAGGGATGTGATCTTACAGAGTAAAGGGAG 1039
QY 241 gaccagatgtagaatacaccagaagaagatagactctttgtgctagtagaagttgctc 300
DB 1040 GACCAGATGTTAGATACACCAAGAGATATCAGCTCTTTGTGCTTAGTAAGTGTGCTC 1099
QY 301 ctgctccgagcgtgtgtgtctcccaaggtctcggaaatgagcaggttcttaccagtg 360
DB 1100 CTGCTCCGGCAGCTGTGTGCTCCACAGGTCTGGAATGACACAGTCTCCACAGGTG 1159
QY 361 tcttcaagatatacccaatcagacattcgtgggtttttagaaggcattatgacct 420
DB 1160 TCTTCACAGTATCCCAATACAGCAATTCGTGGGTATGTGCACAGCGATTATGCAAT 1219
QY 421 caaagaacaacatacctcattatcacttctcactgataatgaggaagatttctgt 480
DB 1220 CAAGAGAAACCATCTCATATTACCTTCTATCATGTAAATATGGGAGAGTATTG 1279
QY 481 tggtaeggaagaacttaataagatatagaaggaaggaagcaaatcttctgtaagact 540
DB 1280 TGTGTACGCAAGAACTTAATTAAGATATTAGCAAGGAGAAACCAAAATTTCTG 1339
QY 541 tcatcataaagctttagcgttggcattttaaagttcccgaaagcaaatcttcttgg 600
DB 1340 TCATCATATAAAGCTTCAGCTTTGGCATGTTTAAAGTCCCGAAGCAAAATCTTCTT 1399
QY 601 tggacacagttataagaacaataatcattgtgtgtagcagtggttgcgtcagtaactc 660
DB 1400 TGGACACAGTTATAGCAAAATCATGTTGTGATGTCAAGTGTTCGCGTACAGTCT 1459
QY 661 caagactcaccacactatgtgtttaaagcacatataaagaagtggaacacattgcta 720
DB 1460 CAGGACTCACACACCTATTGTGTTTAATGCATATTAAGAGAGTGGAACCAATGCT 1519
QY 721 atgatgtgtctctttagcaaccaagaagaagggttaaacacacagccaatgaattcc 780
DB 1520 ATGATGTGTTTCTTTAGCAACCAAGCAAGAGAGGTAAACTACGCCCATGAAATTC 1579
QY 781 agggctgagcacttcaagactcccaattaggaatgtttggaattagaattctctgcta 840
DB 1580 AGGGTGGCATTCTTTACGATCTCCAAATTTAGGAATGTTTGAATTAAGAAATTTCT 1639
QY 841 ttattaacaccactcaagcagctattttggaattgtgtcttagagagtaaacctgtcc 900
DB 1640 TTATTAAACCACTCAAGCAGTATTTTGGCAATGTGCTGTTACAGAGATAAACTGTGC 1699
QY 901 ctgcagataatgaanaagggtttagatgtagcagatgagtcgttaccatcagttgtg 960
DB 1700 CTGCAGATATATGAAAAAGGTTTGAATGTGCTAGCATGATGCTGTATCACTAGTGTG 1759
QY 961 ataacgggtgtgtgtagtagagcagttggaagcccaagtgacct 1001
DB 1760 ATCACCGGTGTGTGATGAGCAGTTGAGCCCAAGTGCTT 1800

```

```

RESULT 2
HSPDCE2 2583 bp mRNA linear PRI 28-Jul-1995
LOCUS 2583 bp mRNA linear PRI 28-Jul-1995
DEFINITION Human mRNA for dihydrolipoamide acetyltransferase (PDC-E2) (EC
2.3.1.12).
ACCESSION Y00978.X13969
VERSION Y00978.1 GI:35359
KEYWORDS dihydrolipoamide acetyltransferase.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2583)
AUTHORS Patel,M.S.
TITLE Direct Submision
JOURNAL Submitted (16-FEB-1989) Patel M.S., Case Western Reserve
University, Dept. of Biochemistry, School of Medicine, 2119
Abington Road, Cleveland, Ohio USA 44106
2 (bases 1 to 2583)
REFERENCE Thekkumkara,T.J., Ho,L., Wexler,L.D., Pons,G., Liu,T.C. and
Patel,M.S.
TITLE Nucleotide sequence of a cDNA for the dihydrolipoamide
acetyltransferase component of human pyruvate dehydrogenase complex
JOURNAL FEBS Lett. 240 (1-2), 45-48 (1988)
MEDLINE 89052894
COMMENT The sequence overlaps with that reported by Coppel et. al. in Proc.
Natl. Acad. Sci. USA 85:7317-7321(1988) J03866. Data kindly
reviewed (29th March 1989) by Patel M.S.

FEATURES
Source
1..2583
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue="type=liver"
/clone.lib="lambda gtl1"
transit_peptide
<1..162
<1..1848
/codon_start=1
/product="PDC-E2 precursor (AA -54 to 561)"
/protein_id="CAA68787.1"
/db_xref="GI:35360"
/db_xref="SWISS-PROT:P10515"
/translation="RVTSRSRGPAPARRRSTVYGGYVALGQMTGPTSSGAPRNRLILO
LLGSGRRYTSIPRHQKPLPSLSTPMAGTIAIMKKKEDGILNGLIAVEFEDKAT
VGFSELECYMAKILVAGTRNDVPIGLATICTVAGPEDEIKFRTYLLDSSAAPTPOAA
PAPTPAATASPTPSAQAGSSYPHMQVILLPALSPYMTMTGVQWKKVGEKLSQD
LIALETETKATIGFEVQDEGYLAKILVEGTRDYPLGPICTIYEKPADISAFADYRP
TEVTLKQVPPPTPPVAAYPTPOPAPATPSAPCAVPATPGKRVFVSLAKLAV
EKGIDLTQVKGSGPDGRITKDDISFVSKAPAPAAVAVPTGPMAPVPTGVDIIP
ISNIRRVIAORIMOSKOTIPIHYLISIDVNMGEVLLVRELMKILEGSKISVNDFIK
ASALACIKVPEANSSMDTVIRONHYVDVSAVSTPGLITIPYFNHITGERTIAND
VSLATKAREGLQPHFQGGFTLSNIGMFGINKFSATIPPACTILATGASBDKLV
PADNKGFDVAASMSVTLSCDHRVVDVAGVQWMLAEFKYLEKPTMLL"
mat_peptide
163..1845
/product="mature PDC-E2 (AA 1-561)"
BASE COUNT 742 a 562 c 564 g 715 t
ORIGIN
Query Match      100.0%; Score 1001; DB 9; Length 2583;
Best Local Similarity 100.0%; Pred. No. 2e-234;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atatacagcatttgctgactaagcccaacgaaglaacagatttaaacacacaagtc 60
DB 800 ATATATACGATTGTGCTGACTATAGGCCAACCAAGTAACAGATTAAACCAAGTGC 859
QY 61 caccactaccacccccggggtggtcgtgttctcccaactcccccagcttagctccta 120
DB 860 CACCACCTACCCACCCCGGGTGGCGGTCTCTCAACTCCCGCCTTAAGTCTCTTA 919
QY 121 caccctcaacacccctgcacagctactccctgagccaaagggaagggtgtgttagcc 180

```

Accession	Gene	Species	Chromosome	Position (kb)	Orientation	Sequence	Length (bp)
AK057299	AK057299	2696 bp	mRNA	linear	PRI 31-OCT-2001		
LOCUS	DEFINITION	Homo sapiens CDNA FLJ13217 f1s, clone TEST12001269, highly similar to DIHYDROLIPOAMIDE ACETYLTRANSFERASE COMPONENT OF PRIVATE DEHYDROGENASE COMPLEX, MITOCHONDRIAL PRECURSOR (EC 2.3.1.12).					
ACCESSION	AK057299	AK057299.1	GI:16552934				
KEYWORDS	oligo capping: f1s (full insert sequence).						
SOURCE	Homo sapiens testis cDNA to mRNA, clone_11b:TEST12 clone:TEST12001269.						
ORGANISM	Homo sapiens						
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.						

REFERENCE AUTHORS	TITLE
1 (sites)	Ishibashi,T., Kanehori,K.,Yosida,M., Matanabe,S., Ishida,S., Ono,Y., Hotuta,T., Hiraoka,S., Murakawa,K., Takiguchi,S., Kusano,J., Matanabe,M., Fujiwara,K., Tanai,H., Ishida,M., Yamasaki,H., Chiba,Y., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Makamitsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Matsuo,K., Nakamura,Y., Sekine,M., Rikuchi,H., Kanda,K., Wagaetsuma,M., Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagahari,K., Masuno,Y., Negai,K. and Isogai,T.
JOURNAL	NEDO human cDNA sequencing project
REFERENCE	Unpublished
AUTHORS	2 (bases 1 to 2696)
TITLE	Isogai,T., Otsuki,T.
JOURNAL	Direct Submission
COMMENT	Submitted (24-OCT-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 297-0812, Japan (E-mail:genomice@hi.co.jp, Tel:81-438-52-3951, Fax:81-438-52-3952) NEO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; CDNA full insert sequencing: Research Association for Biotechnology (RAB); CDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one-pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: RAB and HRI.
FEATURES	Location/Qualifiers
source	1..2696 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="TESTI2001269" /issue_type="testis" /clone_id="TESTI2" /note="Cloning vector: pME18SFU3"
BASE COUNT	654 a 737 c 653 g 652 t
ORIGIN	
Query Match	100.0%; Score 1001; DB 9; Length 2696;
Best Local Similarity	100.0%; Pred. No. 2e-234;
Matches 1001; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY 1	atataacgatttcctgaactataggccaacgcgaagttaacagattaaaccacaagttgc 60
DB 1555	ATATACGACATTTCCTGTGCATAGGCCAACCGAAGTAAACAATTTAAACCACAAAGTGC 1614
QY 61	caccacctaccaccccccgcgggtgcgcgtcgttctctcccaactcccaagccttagtccta 120
DB 1615	CACCACACTACCCCAACCCCCGGGTGGCGGCTTCCTCACAACCTCCCAGCCTTAGTCTCTTA 1674
QY 121	caacttaagcacccctgcgccagctactcctctgtctgacccaagggaagtggttgttaagc 180
DB 1675	CACCTTGACGACCCCTGCCCCAGCTACTCCTGCTGGACCAAAGGAAGGGTGTGTTAGGC 1734
QY 181	clcttgcaaagaagttgacagtagagaagaagattgatcttacacaagaftaaagaagagac 240
DB 1735	CTCTTGCAAGAAGATTGGCAGTAGAGAAAAGGATTGATCTTACACAGTAAAAAGGACAG 1794
QY 241	gaccagatgtagtaaatcaccacaagaagagatatgcactctttgttgcctagtaaagttgctc 300
DB 1795	GACCAGATGGTAGAATCACCAAGAAGGATATGCATCTTTTGTGCTAGTAAAGTTGGTC 1854
QY 301	ctgcctcgaggagctgtttgtcctcccaaggtcctcgtgaatgycacacagttctaagttg 360
DB 1855	CTGCTCCGGCAGCTGTTGTGCTCCACAGGTCTCGGAATGGCACCAAGTCTCTACAGGTG 1914
QY 361	tcttcaagatatacccaatcaagaacattcgcgggttaattgacacgagattaatgcaat 420
DB 1915	TCTTACAGATATGCCAATACGACAACATTTCGTGGGTTATTTCACACGCAATTAAAGCAAT 1974
QY 421	caagcaaaccatcctcaattatcactcttatcgatgtaaatatgygagaagtttgt 480
DB 1975	CAAGCAAAACCATCTCTATTTACCTTTCTATGATGATAATATGTGGAGAAAGTTTGT 2034

BASE COUNT		666 a	644 c	603 g	627 t
ORIGIN					
Query Match	96.0%; Score 961; DB 9; Length 2540;				
Best Local Similarity	98.2%; Pred. No. 1.2e-224;				
Matches	983; Conservative 0; Mismatches 15; Indels 3; Gaps 1;				
Qy	1	atatacgcacatttgctgtaactaaggccaaccgaagtaacagatttaaaacacaaagttgc	60		
Db	1432	ATATATCAGCATTTTGCTGACTATAGGCCAACCGAAGTAAACAATTTTAAACACCAAGTGC	1491		
Qy	61	caccacccaaccccccgcgggtgagccggtgttctccccaactcccaagccttaagtccta	120		
Db	1492	CACCACTTACCCCAACCCCGGGGGCGGTGTCTCCACACTCCCAAGCTTTAGTCTCTTA	1551		
Qy	121	caccttcagcacccctgcgcacagctactcctgtctgagccaaaggaaagggtgttgttaacc	180		
Db	1552	CACCTTGCACACCCCTGCCACGCTACTCTCTCGACACCAAGGAAGGGTGTGTGTGACCC	1611		
Qy	181	ctctctgcaaaagaagttgacgaaglaagaaaggatttgatcttaacaaaglaaaaggagacag	240		
Db	1612	CTCTTGCAAAAGAGTTGGCAGTAGGAAGAAAGGATTGATCTTACACAGTAAAGGGACAG	1671		
Qy	241	gaccagatggtgtagaatcccaagaagagatcagactcttttggctagttaaagttgttc	300		
Db	1672	GACCAAGATGTGATGATATCCACGAAGAAAGATGATGACTCTTTTGTGCTGTAAAGTTGCTC	1731		
Qy	301	ctgctcggcagcgtgtgtgctctcccaagagtcctgtgaatggcaccagttcctaaagtg	360		
Db	1732	CTGCTCCCGCAGCTGTGTGTGCTCTCCACAGGCTCTGAATGGCACACAGTCTCTACAGGTG	1791		
Qy	361	tcttaacagatatcccaatcagcaacatctcgtcaggttatctgcaagagattaagcaat	420		
Db	1792	TCTTACAGATATCCCAATACGCAACATTCGCGGGTATTGACACGCAATTAAGCAT	1851		
Qy	421	caaaagcaaacatatacctaattatcattccttcctatcgatgtaataatggagaagtttgt	480		
Db	1852	CAAAAGCAAAACATACCTCATTTATATACCTTCATGCTGTAAATAT---GGAGAAGTTTGT	1908		
Qy	481	tgttaagaaagaacttaataagatatlaagaaggagaagcaaaattctgtcaatgact	540		
Db	1909	TGTATACGGAAGAACTTATATTAAGATATTGAAGAGGAGCAAGCAAAATTTCTGTCAATGACT	1968		
Qy	541	tcatcataaagaactcagactttgagatgtttaaaggctccgaagcaaatctctcttga	600		
Db	1969	TCAATCATTAAGAGCTTCAGCTTTGGCATGTTTAAAGTTTCCCGAACAATTTCTTTTGA	2028		
Qy	601	tggacacagttataaagaacaaatcatgttgtatgctcagttgttcggctcagtaactcctg	660		
Db	2029	TGGACACAGTATATAAGACAAATATCATGTTGTTGATGTCAAGTGTTCGGCTCAGTACTCTCG	2088		
Qy	661	caggacatataccactatttgtttaaatagcataataaaggagtggaacacattgcta	720		
Db	2089	CAGGACTCATACACCTATTTGTGTAAATGCAATATAAAGAGTGGAAACCAATTTGCTTA	2148		
Qy	721	atgatgtgttcttcttgaacaaccaagaagaagaggttaactacagccaatgattcc	780		
Db	2149	ATGATGTGTTTCTTTAGCAACCAAGCAAGAGAGGTTAACTACAGCCACATGTAATCC	2208		
Qy	781	agggtggcactttaagctctccaatttaagaaagtgttggaaatlaagaattctctgtca	840		
Db	2209	AGGGTGGCACTTTTACGATCTCCAAATTTAGGAATGTTTGAATTAAGAATTTCTCTGCTA	2268		

QY 841 ttattaaaccaccctcaagcatgattttgccaatttgcttcagaggaataaactgttc 900
|||||
Db 2269 TTATTAAACCAACCTCAACGATGATTTTGGCAATGTGCTTCAGAGGATMAACTGCTCC 2328
QY 901 ctgcagaataaagaagggttgatgtagctagcatgactgctctgtacactcaagtgtg 960
|||||
Db 2229 CTGCAGATTAATGAAAAAGGTTGATGTGGCTAGCATGATGTCGTACACTCAGTTGTG 2388
QY 961 atcacgggtggtggtgagagcagtttgagcccaagtgtc 1001
|||||
Db 2389 ATCACCGGTTGTGATGAGACAGTTGACCCAGTGGCTT 2429

RESULT 5
AB036739 2543 bp mRNA linear MAM 03-JUL-2001
LOCUS AB036739
DEFINITION AB036739.1 GI:14587785
ACCESSION AB036739
VERSION AB036739.1
WORDS dihydrolipoamide acetyltransferase; dihydrolipoamide
acetyltransferase.
SOURCE Sus scrofa
ORGANISM Sus scrofa
Mammalia: Eutheria: Chordata: Craniata: Vertebrata: Euteleostomi:
Eukaryota: Metazoa: Chordata: Cetartiodactyla: Suina: Suidae: Sus.
REFERENCE 1 (bases 1 to 2543)
AUTHORS Koike, K.
TITLE Isolation and characterization of the cDNA encoding the
dihydrolipoamide acetyltransferase component of the porcine
pyruvate dehydrogenase complex
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2543)
AUTHORS Koike, K.
TITLE Direct Submission
JOURNAL Submitted (07-JAN-2000) Kichiko Koike, Nagasaki University School
of Medicine, Department of Pathological Biochemistry; Sakamoto
1-12-4, Nagasaki, Nagasaki 852-8523, Japan
(E-mail: kichiko.koike@nms.sakyou.ne.jp, Tel: 81-95-849-7098,
Fax: 81-95-883-5065)

FEATURES
Location/Qualifiers
source 1..2543
/organism="Sus scrofa"
/db_xref="taxon:9823"
CDS 6..1949
/EC_number="2.3.1.12"
/codon_start=1
/product="dihydrolipoamide acetyltransferase"
/protein_id="BAB61720.1"
/db_xref="GI:14587785"
/translation="MMRVCAAPRQNAAPRAGFARWALREBPAGCAAPRAGSVPAR
CSSTTRGGRSRALCMASATRGATPQNRILLQLMGSPNKRMTSLPQHVLPSPSPT
MOAGTJARWEKKEGDKINSEGLAEVETOKATVGFSELECTMAKILAEGRDVPVG
AICITVEKPEDEIAFNKNTLDSAPAPQAAPAPPAAPAPPSAPGSSTPT
MOVLPALSPMTMGTVORMEKVGKESGLAEFDKATIGVEVEGLATIL
IPGRTDVLGTPPLCTIVEKADIPAFVRYRPEVDLAPRAPPSPSVTPVPAPRO
PVAPRPAATRPAPRPAAGKRLVPLAKLAKSGITDLOIGTGSDGRIKKDINSF
VPTKAPPTPAAAVPPSPGVAAVPTGVPFDIPISNRIYAIQRLMSKOTIRYTIYSV
DVNMGVLLVRKELNMLEGRKISVNDTIKASALACLKVPANSSMDTVIRONHV
DIVSAVSTPAGLITPPIVNAHKGLETIANVSLATAREKQLDPHEVOGTFTIS
NLMFGIKNFSAIINPQACILAVGASEDRLEFPADNEKGFVAVMSAVLSCDHRVVD
GAVGQWLAIEFRYLEKPTMLL"

sig_peptide 6..263
mat_peptide 264..1946
polya_signal 1..2145
BASE COUNT 724 a 563 c 576 g 680 t
ORIGIN

Query Match 87.1%; Score 871.4; DB 4; Length 2543;
Best Local Similarity 91.9%; Pred. No. 1..1e-202;
Matches 920; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 1 atatacagatttgcctgactataggccaacgaagtaacagatttaaacccaagtgc 60
|||||
Db 901 ATATACCAGATTTGCTGACTATCGCCAACTGAAGTAATGTATTTAAACCAACGACAC 960
QY 61 caccaccctaccccccccgggttgccgtgtctctccaactcccccccttagctctca 120
|||||
Db 961 CACCACCTTACCCGCTGCCGGGTGACCCCTTCTCTCAGCTGCCCAACCGTAGCCCTTA 1020
QY 121 caccctcagaacctgcgccagctactcctcgtgcacccaagggaaggtgtgttagcc 180
|||||
Db 1021 CACCTCGACGACCCCGCCAGCTACTCTCTGACCAAAAGGAAGGTGTGTGTAGCC 1080
QY 181 ctcttcgaagaaggttggcagtagagaaggtatgactcttaacagaagtaaaaggacag 240
|||||
Db 1081 CTCTTGCAAAAGAAATTTGGCATCGAGAAAGCAATTCATTCACAAATTAAGGGACAG 1140
QY 241 gaccagaatgtagaatcaccagaagaagatctgcactcttttgctcagtaaaagtgtc 300
|||||
Db 1141 GACCAGATGGCAGAAATATATAGAGACATGTACTCTTGTGCTTACTTAAAGCTGCTC 1200
QY 301 ctgctccggcagctgtgtgctctcccaaggctcctggaatgacacagttcctacagtgc 360
|||||
Db 1201 CTACTCCAGCAGCTGCTGTTCTCTCCCGAGCCAGAGAGTGGCAGCAGTTCCACAGGG 1260
QY 361 tcttcaagatattcccaatcaagaacattcgtcggttatgtgaacagcattatgcaat 420
|||||
Db 1261 TCTTCACAGATATCCCAATAGCAACATTCGTGAGTTATTCGACAGCGGTATATCAAT 1320
QY 421 caaagcaaacatcattatcattatcattatcattatcattatcattatcattatcatt 480
|||||
Db 1321 CTAAAGCAAAACCATACCTCATATTACCTTCTGTTGATGATGAATATGAGAAAGTTTGT 1380
QY 481 tggtaacgaaagaactaataagatatagaaggaagaacaaattcctgtaactgact 540
|||||
Db 1381 TGGTACGGAAGAAAGCTTAATTAAGATGTTAGAAAGGAGAACAAATTTCTGTAATGAT 1440
QY 541 tctcataaaaggcttaagcttggcagtgtaaaagttcccggaagaacttctctgtga 600
|||||
Db 1441 TCAATTATTAAGAGCTTCAGCTTGCAATGTTAAAGTTCCTGAAGCAAAATTCCTTTGGC 1500
QY 601 tggacacagttataagaacaaatcattgtgtgtgacagtggttgcgtgtaactcctgc 660
|||||
Db 1501 TGGACACAGTTATTAAGCAAAATCATGTTGTGATATCAATGTCGACAGTACGACCTCG 1560
QY 661 caggactcaccacactattgtgttaatgacacataaaaggatggaacattgtcta 720
|||||
Db 1561 CAGGACTCATCACACCTATTGTTATTAATGCATATTAAGGACCTGGAACCATTTGCTTA 1620
QY 721 atgatgttcttctttagcaacaaagaagaaggtaaacacagccacatgaattcc 780
|||||
Db 1621 ATGATGTTGTTCTTTAGCAACCAAGCAGAGAGGTTAAACTPACACCATGAGAGTCC 1680
QY 781 aggtgtgacattttagacatcccaattttagaaatgtttgaaatlaagaattctctgcta 840
|||||
Db 1681 AGGTTGACATTTTACAAATTCGAATTTAGAGATGTTTGAATTAAGAATCTTCTGCTGA 1740
QY 841 ttattaaaccaccctcaagcatgattttgccaatttgcttcagaggaataaactgttc 900
|||||
Db 1741 TTATTAAACCTCTCTCAACGATGATTTTGGCAATGTGCTTCAGAGATAGACTGTTTC 1800
QY 901 ctgcagaataaagaagggttgatgtagctagcatgactgctctgtacactcaagtgtg 960
|||||
Db 1801 CAGCGGATTAATGAAAAAGGATTTGACCTAGCATGATGTCGTGTTACACTCAGCTGTG 1860
QY 961 atcacgggtggtggtgagagcagtttgagcccaagtgtc 1001
|||||
Db 1861 ATCATGGGTTGTGATGAGACAGTTGACCCAGTGGCTT 1901

RESULT 6
AY044265
LOCUS AY044265 1700 bp mRNA linear ROD 01-NOV-2001

DEFINITION Mus musculus dihydrolipoamide S-acetyltransferase precursor, mRNA,
partial cds.
ACCESSION AY044265
VERSION AY044265.1 GI:16580127
KEYWORDS
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Wang, L., Kaneko, S., Kagaya, M., Ohno, H., Honda, M., and Kobayashi, K.
TITLE Molecular cloning, characterization and expression of
dihydrolipoamide acetyltransferase component of murine pyruvate
dehydrogenase complex in NIH3T3 cell
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1700)
AUTHORS Wang, L., Kaneko, S., Kagaya, M., Ohno, H., Honda, M., and Kobayashi, K.
TITLE Direct Submission
JOURNAL Submitted (09-JUL-2001) First Department of Internal Medicine,
Kanazawa University School of Medicine, Takara Machi 13-1,
Kanazawa, Ishikawa 920-8640, Japan
TURNS Location/Qualifiers
source 1..1700
/organism="Mus musculus"
/db_xref="taxon:10090"
/cell_line="BML-ME"
/tissue_type="hepatoma"
CDS <1..1680
/note="PDC-E2; component of the pyruvate dehydrogenase
complex"
/codon_start=1
/product="dihydrolipoamide S-acetyltransferase precursor"
/protein_id="AAI02400.1"
/db_xref="GI:16580128"
/translation="YISLPPHOKVLPSPISPMOAGTIRKREKBEKISEEDLIAEV
ETDKAIVFESLEECMAKILVEPGRDVPVSIICITVEKQDILFAFNLYLDAIA
AAPQAPAPAPAPAPAPASAPGSYPHMOYLPAISPMTMGTOREKKEVGE
KISEEDLIAETDKATIGFEVDEEGLAKIKHVESTRVPGAPKCIIVERQEDIA
PADYRTEVTSKPOAPAPAPAPVAVPTPOVAPPSAPAPAGRGVYFPLAKKL
AAEKGLDVOVCTGPEGRILIKKIDISFVPSKAPAPAPAPAPAPAPAGVPTD
IPISNRIYIAORLMOISKOTIPHYIISLVVNGEVLVYKELMKLEGGKISVNDFI
IKASALACKLVEANSSMDTVIRONHVVDVAVASTPAGLITPTIVFNHIGLEETIA
SDVSLASAKRGKLOPHFEGGTETISLMGMLINFSAILINPACILTAGSEDK
LIPADNEKGFVDAVSVMSTVLSGDHRVVDGAVGAWMLAEKKYLETML."
mat_peptide 7..1677
BASE COUNT 459 a 414 c 423 g 402 t 2 others
ORIGIN
Query Match 76.8%; Score 768.6; DB 10; Length 1700;
Best Local Similarity 86.4%; Pred. No. 1.5e-177;
Matches 865; Conservative 0; Mismatches 124; Indels 12; Gaps 1;
QY 1 alatalcagaatttgcataatagagcaacgaagtaacagalttaaaacacaagatgc 60
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 644 ATATAGACGATTTCGACGATACAGGCAACAGAGTACAGCTTAAAGCCACAGCAG 703
QY 61 caaacactaccccccccggtggtgcgtcttctccaaatcccccagcctttgacctta 120
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 704 CACCACCTGGCCACCCCGGTGGCAGCTTCTCCACCTCCACCCAGCTGTAGCACCTA 763
QY 121 caacttcagacccctgcgcagctactctctgcagccaaggaagggtgtgtgtgagcc 180
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 764 CCCC-----CTCAGCTGCTCTCTCTGACCAAAAGGAGGTGTGTGTATGTC 811
QY 181 ctcttcgaagaaggttgacagtaagaagaagatgtacttcaacaaaglaaaagggagag 240
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 812 CTCTTGCGAGAGATTGGCAGCAGAAAGGATTGACCTCACCACCAAGTAAAGGAGCG 871
QY 241 gaccagatgttaagatcaccaagaagaagatcactcttttggtgacctagaagtgtctc 300
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 872 GACCAGAAAGGAGAAATCATCAGAAAGGACATGACTCTTTGTGCTCTTCAAGGCTGCTC 931

QY 301 ctgctccgacagctgtttgtgctcccaaggtccctgggaatgcaccagttcctaagtg 360
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 932 CTGCTCCACAGAGTGGCAGTGGCTCCCGCGGTCCAAAGAGTGGCCACAGCTCTCGCAGG 991
QY 361 tcttcagagatccccaacagcaacattcgtccggttatcttgacagagcatatgaacat 420
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 992 TCTTTCACAGACATCCCATCAGCAGCATTCGTGAGTATTCGCGAGAGGCTCATAGCACT 1051
QY 421 caaagcaaacatcaccatcattatcattcattcagatgtaataatggagaagtttgt 480
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1052 CGAAGCAGACTAATACCTCATTTATACCTTCTGTGATGATTAATAGGAGAGGTGCTGT 1111
QY 481 tggtaaggagaagacttaataagatataggaggagaagaacaattctgtcactgact 540
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1112 TGGTACGGAAGCACTTAATTAAGATGCTTGAAGGGAAGGGAAGAAATCTCCGCAACGACT 1171
QY 541 tcatacaaaagctcagcttggtgacgtgttaaaagttcccgaaagcaattctcttga 600
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1172 TCAATTAAAGGCTTACGCTTGGCTGTGTAAGTTCCGAGCAAACTGCTTGA 1231
QY 601 tggacaagattataagaacaataatcattgtgtgagtcagtggttggtcagtaactctg 660
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1232 TGGACACAGTTATTAAGCAAAATCATGTGTGACGTGAGTGTGCTGTGACATACCCCTG 1291
QY 661 cagagctcctacacacctattgttttaagtcacataaaaggagtggaaccattgcta 720
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1292 CAGGACTTATCACCCCTATGTGTTAATGACACACTAAAGAGACTGGAACCATTTGCA 1351
QY 721 atgatgtttctctttagcaacccaagaagaggttaactacacacataatcc 780
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1352 GTGATGTGTGTTCTTAAAGCTCCCAAGCAGAGAGGTTAACTGACCTCATGAGTTTC 1411
QY 781 aggtgtgacatttcacagatcccaattaggaatgtttgaaatgaattctctgcta 840
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1412 AGGTGGAACATTTTAACTCCAACTCCAACTGAGGTGTTGAATTAAGAAATTTCTCTGCA 1471
QY 841 ttaataccacactcaagaatgatttggaattgtgtctcagaaggataaacgtctcc 900
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1472 TTATTTAATCACCCTCAGCGATGATTCGTGCAATCCGCTTCAGAGGTAATCACTGATCC 1531
QY 901 ctgcagataatgaaaaagggtttgtagtgcctagacatgctgtctgaactcaattgtg 960
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1532 CACGACGATATGAGAAAGGCTTGAATGTGATGTGATGTCGCTTACACGACGAGTGTG 1591
QY 961 atcacgggtgtgtgagatgagcaggttgagagccagtgacct 1001
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1592 ATCATCGAGTTGTGATGAGCAGTTGAGCCAGTGGCTT 1632
RESULT 7
RATPDC2 2266 bp mRNA linear ROD 20-APR-2000
LOCUS
DEFINITION Rattus norvegicus mRNA for dihydrolipoamide acetyltransferase.
ACCESSION D10655.1 GI:220837
VERSION
KEYWORDS dihydrolipoamide acetyltransferase.
SOURCE Rattus norvegicus (strain:Sprague-Dawley) cDNA to mRNA,
clone:lib:lambda gtl1.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 2266)
AUTHORS Matuda, S., Nakano, K., Ohta, S., Shimura, M., Yamanaoka, T.,
Nakagawa, S., Tani, K., and Miyata, T.
TITLE Molecular cloning of dihydrolipoamide acetyltransferase of the rat
pyruvate dehydrogenase complex: sequence comparison and
evolutionary relationship to other dihydrolipoamide
acetyltransferases
JOURNAL Biochim. Biophys. Acta 1131 (1), 114-118 (1992)
MEDLINE 92256482
COMMENT These data kindly submitted in computer readable form by: Sadayuki

Db 1025 CAGGACTTATCAACCCCTATTGTTTAAATGACACATAAAGAGACTGGAACCAATTGCTA 1084
Qy 721 atgatgttcttctttagcaagaagaagggttaactaagccacatgattcc 780
Db 1085 GGTGATGTTTCTTCTTACCTCCAAAGCAAGAGAGGTAACCTTCAGCTCAGAGATTCC 1144
Qy 781 aggggtgacctttagcatctcaatttaggaatgtttgaaatgaattctctgcta 840
Db 1145 AGGGTGGGACATTTACATCTCCAACTAGGATGTTCGGAATTAAGATTCTCTCGCA 1204
Qy 841 ttatlaaccacactcaagcatgatttggcaatttggcttcaagagataactggtcc 900
Db 1205 TTATTAACCCACCTCAGCATGATATTGTGCAATTGCTCTCCGAGATTAACCTGATC 1264
Qy 901 ctggcaataaataaagggtttagatgtgcttagcatgattctgtt--tcaactcagttg 958
Db 1265 CTGCAGATTAATGAGAAAGCTTTGACGTGGCTGATGTGATGTCTGCACACACTAGCTG 1324
Qy 959 tgatcacgggtgtgtgtagtgagcagttgagccagtgacct 1001
1325 TGATCATGAGTGTGTGATGAGACAGTGGAGCCAGTGGCTT 1367

RESULT 11
A05169 LOCUS A05169 966 bp DNA linear PAT 04-MAY-1993
DEFINITION Synthetic PRMT 2.2kb cDNA insert of the 70kd mitochondrial antigen of PBC.
ACCESSION A05169
VERSION A05169.1 GI:345029
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 966)
AUTHORS
TITLE PRIMARY BILIARY CIRRHOSIS AUTOANTIGEN
JOURNAL Patent: WO 8804689-A 4 30-JUN-1988;
FEATURES
source Location/Qualifiers
CDS
1..966
/organism="synthetic construct"
/db_xref="taxon:32630"
<1..>966
/codon_start=1
/transl_table=11
/product="Human 70kd mitochondrial antigen of PBC"
/protein_id="CA00431.1"
/db_xref="GI:345030"
/translation="PGSSYPHMQVLLPALSPMTMGTVORNEKKVGEKLSBGLLAE
IFMDKATIGFEVQEGYLAKILVPEGTROVPIGTPLILVEREADISAFADRYFEVT
DLKQVPPPTPPVAAPPTPOPIATPSPAGPAPAGKRGVFPVSLAKLAEKGI
DLQVAGTDPDGRITRKLDIDSYFSPKVAAPAAVPPTPGMAPPTGTFDIPISNI
RVIAORLMQSKOTIPHYLSIDVNGEVLVAKELNKLLEGSKLSVDFITIKASAL
ACLKVEANSMMDTVIRONHVDVSAVSTPAGLITPVPNAHK"

BASE COUNT 289 a 229 c 208 g 240 t
ORIGIN

Query Match 70.0%; Score 701; DB 6; Length 966;
Best Local Similarity 100.0%; Pred. No. 4.9e-161;
Matches 701; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atatatagatttgcgactatagagcaacgaagtaacagattlaaaccacaagtgc 60
Db 266 ATATATAGATTTGCTGACTATAGCCAAACCAAGTAAAGATTTAAACCAAGTGC 325
Qy 61 caccactaccacccccgggtgacgtgttccctcaactccccagccttagctccta 120
Db 326 CACCACTACCAACCCCGGTGGCGCTGCTCCTCAACTCCCACTTTACTCTCTA 385
Qy 121 caactttagcaacctgcccagctactctgtgtaacaaagggaaggtgtttgttagcc 180
Db 386 CACCTTCAGACACCTGCCAGCTACTCTCTGTGACCAAAAGGAGGAGGCTTTGTAGCC 445

Qy 181 cctctgcaagaagttgacagtagagaaggattgattccttacacaagttaaaggacag 240
Db 446 CTCTTGCAAAAGAGTTGGCAGTAGAGAAAGGATTTGATCTTACACACTATAAAGGACAG 505
Qy 241 gaccagatgtagaataccccaagaagatcagactctttgtgcttagtaagtgtc 300
Db 506 GACCAAGATGTAATACCAACCAAGAGATATGCACTCTTTGTGCTGTAAAGTTGCTC 565
Qy 301 ctgtctccggacgtgttgtgctccacaggtcctggaatgacacagttcctcagaatg 360
Db 566 CTGTCTCGGACCTGTGTGTGCTCCACAGGCTCTGGAATGGACCACTTCTTACAGGTG 625
Qy 361 tcttcaacagatacccaatcagaacacattgctcggttatttgcacagcattatgcaat 420
Db 626 TCTTACAGATATCCCAATCAGCAACATTCGTGGGTTATTTCACAGCATTAATATGAAT 685
Qy 421 caaagcaaaccttacctatcttattaccttctcatagatgtaataatggagaagttgt 480
Db 686 CAAAGCAAAACCAATCTCATTTATACCTTTCTATCGATCTAAATATGGAGCAATTTGT 745
Qy 481 tggtaagaaagaaacttaataagataatagaaggagaagcaaaattctgtcaatgact 540
Db 746 TGGTACGGAAGAAGACTTATATAGATATTAGAGGAGGAACAAATTTCTGTCAATGACT 805
Qy 541 tcatcataaagcttcaagctttagcatgtttaaagttcccgaaagcaattcttcttga 600
Db 806 TCATCTAATAAAGCTTTCAGCTTTGGCATGTTTAAAGTTCCCAACCAATTTCTTGA 865
Qy 601 tggacacagttataagaacaataatggtgtttagatgtagcagttgtcggtcagttacctg 660
Db 866 TGACACACTTTATAGACAAATAATCATGTTGTTGATGTCTAGTGTTCGGGTCACTCTGT 925
Qy 661 caggactcatcacacctattgtgtttaaagcacatataaa 701
Db 926 CAGGACTCATACACCTATTGTTTAAATGACACATATAAA 966

RESULT 12
HSDHACTYL
LOCUS HSDHACTYL 660 bp mRNA linear PRI 10-JAN-1994
DEFINITION H.sapiens mRNA for dihydrolipoamide S-acetyltransferase.
ACCESSION X13822
VERSION X13822.1 GI:30523
KEYWORDS
acetyltransferase; dihydrolipoamide S-acetyltransferase; lipoate
acetyltransferase; pyruvate dehydrogenase.
human.
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 (bases 1 to 660)
Moehario, L.H.
TITLE Direct Submission
JOURNAL Submitted (22-DEC-1988) Moehario L.H., Department of Biochemistry
and Centre for Molecular Biology and Medicine, Monash University,
Clayton, Victoria 3168, Australia
2 (bases 1 to 660)
Moehario, L.H., Smoother, P.M., Devenish, R.J., Mackay, I.R.,
Gershwin, M.E. and Matzki, S.
Nucleotide sequence of a cDNA encoding the lipoate acetyl
transferase (E2) of human heart pyruvate dehydrogenase complex
differs from that of human placenta
Biochem. Int. 20 (2), 417-422 (1990)
MEDLINE 90197673
COMMENT Related sequence: J03866
FEATURES
source Location/Qualifiers
1..660
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="heart"
/clone_lib="cDNA library in lambda gt11"
<1..>660
/EC_number="2.3.1.12"


```

Db      252 GTTACGTGATGTCCTGTTACCTCAGCTGTGATCATCCAGTTGTGATGAGACGATTGGA 311
Oy      990 gccacgtgctt 1001
        |||||
Db      312 GCCCAGTGCTT 323

RESULT 14
AX305911      319 bp      DNA      linear      PAT 11-DEC-2001
LOCUS         AX305911
DEFINITION    Sequence 662 from Patent WO018188.
ACCESSION     AX305911
VERSION       AX305911.1 GI:17645279
KEYWORDS
SOURCE        house mouse.
ORGANISM      Mus musculus
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE     1
AUTHORS       Ishikawa,K., Asai,S., Takahashi,Y., Nagata,T. and Ishii,Y.
TITLE         Method for examining ischemic conditions
JOURNAL       Patent: WO 018188-A 662 22-NOV-2001.
FEATURES
             1. 319
             location/Qualifiers
BASE COUNT    89 a      60 c      77 g      93 t
ORIGIN

Query Match      22.3%; Score 223; DB 6; Length 319;
Best Local Similarity 88.9%; Pred. No. 4.3e-44;
Matches 241; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

Oy      731 ttcttagcaaccaagaagaggttaactacagccatgtaattccaggttgccac 790
        |||||
Db      1  TTCTTTAGCCTCCAAAGCGAGAGGGTAAACTTCAGCCCTCAGAGTTCAGAGGTGAAAC 60

Oy      791 tttaagatcccaattgaagaatgtttgaataagaattctctcgtacttaacc 850
        |||||
Db      61 ATTTCACATCTCCACTTACGAGATGTTTGCATTAAGAAATTTCTGCAATTAGTAGTGC 120

Oy      851 acctcaagcatgatttttgcaatttggtctcaagaagataaacctggtccctgcagataa 910
        |||||
Db      121 ACCTCAGCGCATGTATTCTGCAATCGGTCTCATAGGATMAAATGATCCACAGATTA 180

Oy      911 tgsaaaagggttgatggtgctagcagcatgctgtgttaacactcaagttgataccaggt 970
        |||||
Db      181 TGAGAAAGGCTTGATGTGCTGATGTGATGTGCTGATGCTGATGCTGATGCTGATGCTGATG 240

Oy      971 gctgagatgagacagttgagagccacgtgctt 1001
        |||||
Db      241 TGTGATGAGAGCAGTTGAGACCCAGTGCCTT 271

RESULT 15
DD006634      1833 bp      mRNA      linear      INV 08-MAR-1994
LOCUS         DD006634
DEFINITION    Dictyostelium discoideum dihydrolipoamide transacetylase mRNA,
partial cds.
ACCESSION     DD006634
VERSION       DD006634.1 GI:458425
KEYWORDS
SOURCE        Dictyostelium discoideum.
ORGANISM      Dictyostelium discoideum
               Eukaryota; Mycetozoa; Dictyostelidia; Dictyostelium.
REFERENCE     1
AUTHORS       Mueller-Taubenberger,A.
TITLE         Dihydrolipoamide transacetylase gene from Dictyostelium discoideum
JOURNAL       Unpublished
REFERENCE     2 (bases 1 to 1833)

```

```

AUTHORS       Mueller-Taubenberger,A.
TITLE         Direct Submission
JOURNAL       Submitted (08-FEB-1994) Annette Mueller-Taubenberger,
               Max-Planck-Institute for Biochemistry, Am Klopferspitz 18a,
               Martinsried, 82152, Germany
FEATURES
             1. 1833
             location/Qualifiers
SOURCE        Dictyostelium discoideum
               /organism="Dictyostelium discoideum"
               /strain="AX2"
               /db_xref="taxon:44689"
               /clone_id="lambda-gc11"
               <1..1779
               /note="subunit E2 of the dehydrogenase complex; product is
               located in the mitochondrion; putative single copy gene,
               but 2 mRNAs of 1.5kb and 2.1kb expressed"
               /codon_start=1
               /product="dihydrolipoamide acetyltransferase"
               /protein_id="AA16511.1"
               /db_xref="GI:458426"
               /translation="FNNOTPKIFTSNNVLFSFSSSSNVFSEILNKRYSKGEI
               TMPALSPMTVGNIVOMKKEGDOIKADVIDREYEDKATWDSYEDNGYLAKILIE
               GTKGEIKPIAIVYKREDESAVKRPPSSASSTPVEAPKPROEAPKSTKY
               PAHRYVGMPLSPSMENGIASWTKKEDQIKADIAVEYEDKATWDSYEDNGYL
               AKILVPGSTSGIQTNPVCIIVKNEKEDKRDYDVEQSSSSSSQESPPSSSS
               SQESTPSSSSQDQYTKKSGERIFATPAKFEASSGIDLSINGTGNRILKADYLE
               FVPOKEVAAQDQDQDTTTRKPTPTSGEETDIPHSNIRKVTAAKTESKOTIPH
               YLTMECRVDKILKRLSELNMTVKISYNDIVKASLPALNDNVNSTWDFIRRY
               HNIDINAVNTPOGLETPIVRGVDKSNISISYKOLAFAONGKLPSEFSEGTFT
               ISNIGMIGIKOFAAVINPPOAIIALVPOKLVFSNNPDSPEFATILSVTLCSDHR
               VVIDAVGAEMLKSFNDYENPKIL"
BASE COUNT    701 a      347 c      276 g      509 t
ORIGIN

Query Match      21.0%; Score 210; DB 3; Length 1833;
Best Local Similarity 54.8%; Pred. No. 7.9e-41;
Matches 489; Conservative 0; Mismatches 385; Indels 18; Gaps 3;

Oy      121 caecttcagcaccctccagctactccctgctggaacaaaggaaggttgtagcc 180
        |||
Db      845 CATCACATCATATCATCAACAACAACATACAGCTAAATCTGGTGAAAGAAATTTGCTGACAC 904

Oy      181 ctcttgaagaagcttgagcagtagaagaaggttgctctcaacaagtaaaaggaagcag 240
        |||
Db      905 CAGCAGCAAGATTTGAGGCACTCATCAAAAGATATGATCTTTACAGCAATTAAGTAGTAC 964

Oy      241 gaccagatgtagaataccacaagaagatacgactcttctgctagtaaaagttgctc 300
        |||
Db      965 GTCCAAATATATAGAAATCTTAAAGCTGATGTTTAAATTTGACCCCAAAAACAAGAG 1024

Oy      301 ctgctccgagagctgtgtgctc-----ctccacaaggtccctggaatgacacagctc 351
        |||
Db      1025 TTGCTCAACAACAACAACAACAACAACACTACTACACCTTAAAAAACCAACTACCAACTT 1084

Oy      352 ctacagtgcttcacagatatacccaatcagcaacatcgctcggtatattgacacagcat 411
        |||
Db      1085 CATCAGGTGAATTCACATGATATATATATATATATATATATATATATATATATATATATATAT 1144

Oy      412 taatgcaatcaagaacacatacctatattactcttctatcagatgtaaatatgagag 471
        |||
Db      1145 TTACAGAAATCAAAACAACATTCACATTAATATTCACAAATGGAATGTGTTGATATA 1204

```

```
OY 472 aagtttggtaggaagaagacttaataagatatagaagggaagaacaaatttctg 531
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1205 AACTCTTAAACTTCTGTCAGATTAAGCAT-----GAACACTGTTAAATCTCTG 1258
OY 532 tcaatgacttcatcaataaagcttcagcttggcatgtttaaaagtcocgaagcaatt 591
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1259 TCAATGATTTCATTGTTAAAGATGCGCGCCCTTCGTGATTAATCCAGTCGTAANT 1318
OY 592 ctcttggatgagacaagtlataagaacaaatcatgttctgtatgtcagtgctgca 651
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1319 CAACATGGACCGCATCAATTCATTAGAGATACACACATTCGATATCAACGTTGCCGTCA 1378
OY 652 gtacccctggcagactcatcacacctatgtttaaagcacataaaaggagtgaaa 711
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1379 ATACACCCACAAAGTTTATTCACTCCATCGTTAGAGGTGTCGATATGAAGGTCCTCAACT 1438
OY 712 ccattgctaattgattgttctcttagcaaccacaagagaagggttaactacagccac 771
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
    1439 CTATCTCAAACTCTGTTAACAATTTGGCTGAANAAGCTCAAAAATGTAAATTAACCCAT 1498
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 772 atgaattccagggtggcaactttagcatcccaatttaggaatgttggaaattaagaatt 831
    ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |
    1499 CCGAATTGAAAGTGTACTTCACTATCTTAACCTTGGGTATGCTCGTATTAAACAAAT 1558
OY 832 tcttgcattatlaaccacacctcaagcatgtatttggcaatttggcttcagaggata 891
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1559 TCGCAGCACTCATCAATCCACACCAAGCTGCTATCTTGCGCTTGTTACACAGAAACTCG 1618
OY 892 aactggtccclgcagataa---tgaaaaagggttggatgtgcttagcatgaltcgtta 948
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
    1619 TGTGTTTTTAAGCAATAAACCAAGACTCACCATATGAACCGCTACATTTTATCCGTTA 1678
OY 949 caactcagtttgatcacccgggtgtgtgagtgagagcaagtcgagccagtgct 1000
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
    1679 CCTTAAGTTGATCATGCTGTATCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1730
```

Search completed: August 13, 2002, 19:30:33
Job time: 6865 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 13, 2002, 18:26:23 ; Search time 208.64 Seconds

(Without alignments)
8237.305 Million cell updates/sec

Title: us-09-761-580-1_COPY_800_1800

Perfect score: 1001

Sequence: 1 atatacagattgctgac.....cagltgagccagtgctt 1001

Scoring table: IDENTITY_NUC
Gap 10.0, Gapext 1.0
1736436 seqs, 858457221 residues

al number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: N_Geneseq_032802:*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1001	100.0	2583	22	AAD07345 Human pyruvate dehydrogenase complex E2 (PDC E2) gene.
2	1000.2	99.9	3859	21	AAE21916 Human breast and o
3	735.8	73.5	1370	9	AAH82410 Plasmid pRRIT enco
4	699.4	69.9	966	9	AAH82411 Insert from cDNA p
5	250.4	25.0	360	23	AAH70149 DNA encoding novel
6	223	22.3	319	24	AAH98637 Mouse ischaemic co
7	209.8	21.0	1502	23	ABL20587 Drosophila melanog
8	209.8	21.0	1850	23	ABL09105 Drosophila melanog
9	155.4	15.5	639	23	AAH70150 DNA encoding novel

10	125.8	12.6	4145	23	ABL20586 Drosophila melanog
11	125.8	12.6	4983	23	ABL09104 Drosophila melanog
12	125.6	12.5	2348	21	AAH46977 Arabidopsis thalia
13	96.6	9.7	2595	23	AAH92326 DNA encoding novel
14	88.2	8.8	22863	21	AAH81500 N. meningitidis pa
15	88.2	8.8	34980	21	AAH21609 Neisseria meningit
16	88	8.8	1302	22	AAH52969 S. epidermidis ope
17	88	8.8	3133	22	AAH54299 S. epidermidis gen
18	88	8.8	3190	22	AAH54272 S. epidermidis gen
19	88	8.8	3354	22	AAH54399 S. epidermidis gen
20	88	8.8	3501	22	AAH53982 S. epidermidis gen
21	85	8.5	1549	21	AAH44948 Arabidopsis thalia
22	85	8.5	1708	22	AAH06839 Arabidopsis pyruva
23	81.8	8.2	4858	18	AAH74690 Staphylococcus aur
24	80.8	8.1	1971	23	AAH80749 DNA encoding novel
25	80.4	8.0	580073	18	AAH58840 Mycoplasma genital
26	79.4	7.9	1290	23	AAH52165 Staphylococcus aur
27	79.4	7.9	1293	23	AAH55200 Staphylococcus aur
28	78.8	7.9	1260	22	AAH53561 S. epidermidis ope
29	78.8	7.9	3394	22	AAH54265 S. epidermidis gen
30	77.8	7.8	482	21	AAH35785 Zea mays DNA fragm
31	77.6	7.8	1230	23	AAH53536 Haemophilus influe
32	77.6	7.8	119211	22	AAH28553 Genomic fragment #
33	77.6	7.8	1830121	17	AAH42063 Haemophilus influe
34	76	7.6	1044	23	AAH55575 Streptococcus pneu
35	76	7.6	11340	19	AAH52280 Streptococcus pneu
36	75.8	7.6	7953	18	AAH74463 Staphylococcus aur
37	74.6	7.5	269223	22	AAH28554 Genomic fragment #
38	72.4	7.2	1254	20	AAH07106 Staphylococcus aur
39	72.4	7.2	1254	20	AAH07107 Staphylococcus aur
40	72.4	7.2	1254	20	AAH07108 Staphylococcus aur
41	71.8	7.2	1702	21	AAH35957 Arabidopsis thalia
42	71.8	7.2	1839	21	AAH38982 Arabidopsis thalia
43	70.2	7.0	1038602	20	AAH01425 Complete genome se
44	69.8	7.0	1185	22	AAH52579 S. epidermidis ope
45	69.8	7.0	2706	20	AAH81790 Granulocytic Ehrli

ALIGNMENTS

RESULT 1	
AAD07345	standard: DNA; 2583 Bp.
ID	AAD07345
XX	12-SEP-2001 (first entry)
AC	Human pyruvate dehydrogenase complex E2 (PDC E2 or PDC E2) gene.
XX	Human: pyruvate dehydrogenase complex E2; PDC E2; polymorphism;
KW	lactic acidemia; coronary failure; cardiac myopathy; weakness; diabetes;
KW	muscle ataxia; peripheral vascular disease; myocardial ischemia; asthma;
KW	bioinformatic analysis; pharmacogenetic analysis; obesity; sepsis; drug;
KW	genetic marker; therapy; ds.
XX	Homo sapiens.
OS	
XX	
FX	Location/Qualifiers
FT	211..1848
FT	/tag= a
FT	/product= "Human pyruvate dehydrogenase complex E2"
FT	replace (857, C)
FT	/tag= C
FT	/note= "This variation occurs during polymorphism"
FT	variation
FT	/replace (1255, A)
FT	/tag= d
FT	/note= "This variation occurs during polymorphism"
XX	EP118679-A1.
PN	
XX	
PD	25-JUL-2001.

PR New human breast and ovarian cancer associated gene sequences and the
 PR polypeptides encoded by these genes, useful in the prevention,
 PR treatment and diagnosis of cancer, immune disorders, cardiovascular
 PR disorders and neurological diseases.

PS Claim 1; Page 715-716; 1299pp; English.

CC Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
 CC proteins AAB56711 - AAB59128. The DNA and protein sequences are
 CC associated with breast and ovarian cancer. Included in the invention are
 CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the
 CC isolation and characterisation of the DNA and protein sequences of the
 CC invention. The breast and ovarian cancer associated DNA, protein, agonist
 CC or antagonist sequences exhibit cytostatic; immunosuppressive;
 CC or neuroprotective; antiviral; antiallergic; hepatotropic;
 CC antidiabetic; antiinflammatory; antitumor; vulnerary; anticonvulsant;
 CC antibacterial; antitumor; antiparasitic and cardiac activity. The
 CC polynucleotide and protein sequences are used in the diagnosis of cancer,
 CC particularly breast and ovarian cancer. The nucleic acid sequences,
 CC proteins, agonists and agonists may also be used in the diagnosis,
 CC prevention and treatment of immune disorders e.g. Addison's disease,
 CC allergies, autoimmune hemolytic anaemia, autoimmune thyroiditis,
 CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
 CC arthritis and ulcerative colitis; cardiovascular disorders such as
 CC myocardial ischaemias; wound healing; neurological diseases such as
 CC cerebral anoxia and epilepsy; and infectious diseases.

CC Sequence 3859 BP; 1162 A; 760 C; 820 G; 1106 T; 11 other;

Query Match 99.9%; Score 1000.2; DB 21; Length 3859;

Best Local Similarity 99.8%; Pred. No. 3.3e-278;

Matches 999; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 atatatcagcattgttgctataggcaacgaagtaacagatttaaacacaaagtc 60
 DB 1069 atatatcagcattgttgctataggcaacgaagtaacagatttaaacacaaagtc 1128
 QY 61 caccactaccacaccccggtgctgctgtctctcctcaactcccgagcttagctctca 120
 DB 1129 caccactaccacaccccggtgctgctgtctctcctcaactcccgagcttagctctca 1188
 QY 121 cactctcagcaccctcctcagcactcctgctgaccaaagggaaggtgttgtagcc 180
 DB 1189 cactctcagcaccctcctcagcactcctgctgaccaaagggaaggtgttgtagcc 1248
 QY 181 cctctgcaagaagttgtagcagtagaagaaggatgattcttacaagaftaaaggagcag 240
 DB 1249 cctctgcaagaagttgtagcagtagaagaaggatgattcttacaagaftaaaggagcag 1308
 QY 241 gaccagatgtagaatacaccagaagaagatagcattcttggcttagtaagttgctc 300
 DB 1309 gaccagatgtagaatacaccagaagaagatagcattcttggcttagtaagttgctc 1368
 QY 301 ctgctcggcagctgtgtgctccacaggtctcgtgaatgagcagattcctcacaagt 360
 DB 1369 ctgctcggcagctgtgtgctccacaggtctcgtgaatgagcagattcctcacaagt 1428
 QY 361 tcttcacagatattcccaatcagcaactcgtcgggtatttgcacagcgattatgcaat 420
 DB 1429 tcttcacagatattcccaatcagcaactcgtcgggtatttgcacagcgattatgcaat 1488
 QY 421 caaagaacacacattatattacattctcctatgcatgtaaatatgaggaagtttgt 480
 DB 1489 caaagaacacacattatattacattctcctatgcatgtaaatatgaggaagtttgt 1548
 QY 481 tggtagaagaagaacttaataagataltgaagaaggagaagaacattctgtcaatgact 540
 DB 1549 tggtagaagaagaacttaataagataltgaagaaggagaagaacattctgtcaatgact 1608
 QY 541 tcatcataaagctcagcttggcatgtttaaagttcccggaagaacattctcttggga 600
 DB 1609 tcatcataaagctcagcttggcatgtttaaagttcccggaagaacattctcttggga 1668

QY 601 tggacacagttataagaacaacatcattgttgatgctagctgttgcggtcagttactctg 660
 DB 1669 tggacacagttataagaacaacatcattgttgatgctagctgttgcggtcagttactctg 1728
 QY 661 cagagctacacacacattgtgtttaatgacacataaaaggagtggaacacattgcta 720
 DB 1729 cagagctacacacacattgtgtttaatgacacataaaaggagtggaacacattgcta 1788
 QY 721 atgagttgttctctttagaacaacaaaggaggttaacttaacgcacattgattcc 780
 DB 1789 atgagttgttctctttagaacaacaaaggaggttaacttaacgcacattgattcc 1848
 QY 781 aaggtgacatttcaacatcccaattagaaatgtttgaaatlaagaattctctgcta 840
 DB 1849 aaggtgacatttcaacatcccaattagaaatgtttgaaatlaagaattctctgcta 1908
 QY 841 ttattaacaccctcaacagatgatttggcaattgtgcttccagagataactggtcc 900
 DB 1909 ttattaacaccctcaacagatgatttggcaattgtgcttccagagataactggtcc 1968
 QY 901 ctgacagataaagaagggtttagatgtgctgtagcattgtctgttaccactcagttgtg 960
 DB 1969 ctgacagataaagaagggtttagatgtgctgtagcattgtctgttaccactcagttgtg 2028
 QY 961 atcaccgggtggtgtagatgagcagttgagccagttgctt 1001
 DB 2029 atcaccgggtggtgtagatgagcagttgagccagttgctt 2069

RESULT 3
 AAN82410
 ID AAN82410 standard; DNA; 1370 BP.
 XX
 AC AAN82410;
 XX
 XX 23-NOV-1990 (first entry)
 DE Plasmid pRMIT encoding 70 kD mitochondrial autoantigen of primary
 DE biliary cirrhosis.
 XX
 XX Primary biliary cirrhosis autoantigen; ss.
 OS Rattus rattus.
 XX
 PN W08804689-A.
 XX
 PD 30-JUN-1988.
 XX
 PF 16-DEC-1987; 87MO-AU00427.
 XX
 PR 16-DEC-1986; 86AU-0009523.
 XX
 PA (AMRA-) AMRAD CORP LTD.
 PA (COPP/) COPPEL RL.
 PI Coppel RL, Gershwin ME;
 XX
 XX WPI, 1988-190632/27.
 DR P-PSDB; AAB82931.
 XX
 PT Novel DNA encoding primary biliary cirrhosis autoantigen - used to
 PT express polypeptide useful in diagnostic tests and for treatment of
 PT prim biliary cirrhosis.
 XX
 PS Claim 1; Fig 6; 42pp; English.
 XX
 CC The clone was isolated from a rat liver cDNA lambda gt 11-amp3
 CC library. It expresses an antigen which is specifically reactive
 CC with sera from patients with primary biliary cirrhosis (PBC). The
 CC clone does not contain the full-length sequence of the antigen. The
 CC sequence is not present in mitochondrial DNA; the 70 kD protein is
 CC therefore coded for by nuclear genes. The protein can be used in

CC highly sensitive ELISA or to treat PBC.
CC See also AAN82411.
XX
SQ Sequence 1370 BP; 368 A; 349 C; 334 G; 319 T; 0 other

Query Match	73.5%;	Score 735.8;	DB 9;	Length 1370;
Best Local Similarity	85.6%;	Pred. No. 4.5e-202;		
Matches 859;	Conservative 0;	Mismatches 127;	Indels 17;	Gaps 3.

QY	1	atattcaagcattcttgctgatactagcccaacccgaagtaacagatttaaaacacaaagctgc	60
DB	380	atatagcacatttgtagagatacgaagcccaacgaagtgcacagcttaaaagccaaagccac	439
OY	61	caaccacctaaccaccccccgctgagccgctgtctctccaactcccacgaccttaagctcta	120
DB	440	caacacactctccacccccccagctgagcagctgtctctctccctccctcccccagcctttaagcactta	499
OY	121	caacttcagcaacccttgcccgagctacccctgcctcgagaccaaaggaaggggtttgttagcc	180
DB	500	cccc-----ctcagcgcgtctcgcctgagccaaagggaaggggttcgttagcc	547
OY	181	ctcttcgcaagaagctgtagagtagaagaggattatcttaacaagaagtaaaagggagag	240
DB	548	ctcttgcaagaagaattgtagcagcagaagaaggattatcctcccccacagttaaaggagacg	607
OY	241	gaccagatggttgatgacacacaaagaagatagcagccttttggcttaagattgctgc	300
DB	608	gaccagaagaagcagaacacacaaagaagacaatggaacctttgtgcttaagtagctgcgc	667
OY	301	ctgctccgagcagctgtgttgctccccaaggtctcggaaatgagcaccagttcctacaagtgc	360
DB	668	ctgccgctgagagagctg--ctcccccggctccaaagagtgcgacacaaactccgcagctgc	724
OY	361	tccttcacatatcccaatccagcaacattcgtccggctttatggcacagcagattaaagcaat	420
DB	725	tccttcatagaacccccacacgaacaaatccgcagatgattgcgcagaggtctcatgacgt	784
OY	421	caaaagaacacatacctctatattacctttctatcgaatgaatatgaggaagatttgt	480
DB	785	cgaagcagactatacctctattaccttctctgttatgtaaatgaggaaggtgctgc	844
OY	481	tgttaaggaagaagacttaataagatttagaagaagggaagcaaaatttcgttaagact	540
DB	845	tgtgtggaaggaagcacttaataaagaatgcttgcgaaggtcaagaaaactccgcgtcaagact	904
OY	541	tcgtcataaagaagcttaagcttgcacatgtttaaagttcccggaagcaaatcttctctgga	600
DB	905	tcctcataaagaagcttcagcttgcgtcctggaaggttcctggaagcaaacctcatcttgcga	964
OY	601	tggacacagttataagacaaataacatcagtgtgtatgcagtgtgtgcggtcagtaacccctg	660
DB	965	tggacacagttataagacaaataacatcagtgtgtatgcagtgtgtgcggtcagtaacccctg	1024
OY	661	caggaactacacacacctatgtgtttaatgacatatlaaagaaggtggaacacctgtgta	720
DB	1025	caggaactatcacacctatgtgtttaatgacacataaagaagatcggaaacatgtgta	1084
OY	721	atgagtgtgtcttcttaagcaaccaaagcaagagaggtlaaactacagccacatgatactc	780
DB	1085	gtgattgtgtcttcttaagctccaaagcaagagaggtlaaacttcagccacagagttgc	1144
OY	781	agggtggaacttttagacatccaatattaagaagtgtttggaattagaataattctctgcta	840
DB	1145	agggtggaacttttagacatccaatattaagaagtgtttggaattagaataattctctgcta	1204
OY	841	ttaataaccacacaaagcaltgattatttggcaalttggtctcaagagataaactgtgcc	900
DB	1205	ttaataaccacacacacagcaltgattatttggcaalttggtctccagagataaactgtgcc	1264
OY	901	ctgcagataatgaaaaagggttgatggtgctgcacgaatgctgc--taccatcaggtgc	958
DB	1265	ctgcagataatgaaaaagggttgatggtgctgcagtgagtgatgcgtgtcaacacactcagctg	1324

Qy	959	tgatcacccggtgtgtgatgagcagtttgagcccaagtggctt	1001
Db	1325	tgatcatcgagttgtgatgagcagttcgagcccaagtggctt	1367

RESULT	4
AA082411	
ID	AA082411 standard; DNA; 966 BP.

AC AAN82411;

DT 23-NOV-1990 (first entry)

DE Insert from cDNA plasmid encoding autoantigen of primary biliary

DE
CY

Primary bills

OS Homo sapiens.

PN W08804689-A.
YY

PD 30-JUN-1988.
YY

PF 16-DEC-1987; 87WO-AU00427.
 XY

PR 16-DEC-1986; 86AU-0009523.

PA (AMIRA-) AMRAD CORP LTD.
PA (COPP/) COPPEL RL.

PI Coppel RL, Gershwin ME:

AA
DR WPI: 1988-190632/27.

LN F 50DD, HAF 02334.
XX

11	move2 path encoding path
PT	express polypeptide use

PARTIAL LIST OF CHANGES

XX

3	0434m	+	+	9	0	434P1
XX						

CC library. It was obt'd. h

sera from patients with

chondrial antibodies on

XX

Query Match	69
-------------	----

Matches 700; Conservati

Qy	1
atatatcagcattgctg	

Db 266 atatcagcattgctg

61 caccactaccaccacc

Db 326 caccactaccacca

Qy 121 caccctcagcaccctgc

Db 386 caccctcagcaccctgc

QY 181 ctctgcaagaagtgc

Db 446 ctctgcaagaactg

QY 241 gaccagatgtagaattaccacgaagagatagcagcttttgccttagtaagttgctc 300
 |||||||
 DB 506 gaccagatgtagaattaccacgaagagatagcagcttttgccttagtaagttgctc 565
 QY 301 ctgctccgagcagctgttgcctccacagagtcctggaatggcaccagttcctacagtg 360
 |||||||
 DB 566 ctgctccgagcagctgttgcctccacagagtcctggaatggcaccagttcctacagtg 625
 QY 361 tcttcacagatattcccaatcagcaaatctgcgcggtatttgcacagcgattatgcaat 420
 |||||||
 DB 626 tcttcacagatattcccaatcagcaaatctgcgcggtatttgcacagcgattatgcaat 685
 QY 421 caaagaacacatcattcattatccttctatcgaatgaaatgaggaagatttctg 480
 |||||||
 DB 686 caaagaacacatcattcattatccttctatcgaatgaaatgaggaagatttctg 745
 QY 481 tggtagcgaagaagacttaataagatatagaagaggaagcaaaattctgtcaatgact 540
 |||||||
 DB 746 tggtagcgaagaagacttaataagatatagaagaggaagcaaaattctgtcaatgact 805
 QY 541 tcatcataaagcttcagcttgcgcatgtttaaagttcccggaagcaaatcttcttga 600
 |||||||
 DB 806 tcatcataaagcttcagcttgcgcatgtttaaagttcccggaagcaaatcttcttga 865
 QY 601 tggacacagttataagacaanaatcattgttgcagtgatgcagtgctcagtaactctg 660
 |||||||
 DB 866 tggacacagttataagacaanaatcattgttgcagtgatgcagtgctcagtaactctg 925
 QY 661 cagagctcatcacactattgttgaatgacataaa 701
 |||||||
 DB 926 cagagctcatcacactattgttgaatgacataaa 966

RESULT 5

AAST0149
 ID AAST0149 standard; cDNA; 360 BP.

AC AAST0149;
 DT 13-FEB-2002 (first entry)
 DE DNA encoding novel human diagnostic protein #5953.
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW Food supplement; medical imaging; diagnostic; genetic disorder; ss.
 YX Homo sapiens.

F# WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR WPI: 2001-639362/73.

DR P-PDB; ABG05962.

PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -

PS Claim 1; SEQ ID No 5953; 103bp; English.

XX The invention relates to isolated polynucleotide (I) and

CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 360 BP; 106 A; 97 C; 81 G; 76 T; 0 other;

Query Match 25.0%; Score 250.4; DB 23; Length 360;
 Best Local Similarity 99.6%; Pred. No. 3,6e-62;

Matches 251; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 atatacagcattgtgcattatagccacgaagtaacagatttaaacacacagtcg 60
 |||||||
 DB 109 atatacagcattgtgcattatagccacgaagtaacagatttaaacacacagtcg 168
 QY 61 caccacccaccaccccccggcgtgttccccaactcccgaccttagctccta 120
 |||||||
 DB 169 caccacccaccaccccccggcgtgttccccaactcccgaccttagctccta 228
 QY 121 caccctcagacccctgcgcagctactcctcttgacacaaagggaagggtgtttagcc 180
 |||||||
 DB 229 caccctcagacccctgcgcagctactcctcttgacacaaagggaagggtgtttagcc 288
 QY 181 ctcttgcaagaagttgacagtagaagaaggatgattcttaacacaagtaaaaggacag 240
 |||||||
 DB 289 ctcttgcaagaagttgacagtagaagaaggatgattcttaacacaagtaaaaggacag 348
 QY 241 gaccagatgta 252
 |||||||
 DB 349 gaccagatgta 360

RESULT 6

ID ABI99637 standard; cDNA; 319 BP.

AC ABI99637;

DT 07-MAR-2002 (first entry)

DE Mouse ischaemic condition related cDNA sequence SEQ ID NO:662.

KW Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
 KW vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.

OS Mus musculus.

PN WO200188188-A2.

PD 22-NOV-2001.

PF 18-MAY-2001; 2001WO-JP04192.

PR 18-MAY-2000; 2000JP-0145977.

XX (UYN1-) UNIV NIHON SCHOOL JURIDICAL PERSON.

QY	573	aaagttccccaagcacaatcttctcttgatgagcacacagttatagaacaaatcatgttgt	632
Db	1023	aaggttccccaagcgaactccgctggtatgacacaagtatttcgaagtagcagcagtc	1082
QY	633	gatgtcagttgttcggtcagtaactcctctcagagactcatcacacctatttgtttaatga	692
Db	1083	gatgttctcggtgttcgtctccacagacaaggctctgattaccgccgattgttttcaatgcc	1142
QY	693	catataaagaagatggaaccattcctaattgattgtttcttcttagcaaccacaaga	752
Db	1143	gaccgcaaaagggtgttcctgtgagatctcttaagatgtcaaggcgctgtgcagccaaggcgccgc	1202
QY	753	gagggtaaacacacacacacatgaattccagggtgagcatttcaagatcccaattcaaga	812
Db	1203	gacacacaaactcagccccacagaattccaggtgtgagaccatctcgtttctaactcgggc	1262
QY	813	atgtttggaattaaagaatttctctgtctataattaaccaactcaagcatgtattttgca	872
-	1263	atgttcggtgtgttaacacagtttgcgctgtcatalcaaacctctcaatcgtgcacctctgc	1322
Db	873	attgtgtcttcagagataaactgttccctg---cagataatgaaaagggtttgattg	929
Db	1323	atggtgcaccacacaagaaacagttgtgtctgtatcccgacagctcgaaggcttcaaggag	1382
QY	930	gtctaacagatgtctctgttaacactcagttgtgatcaccgggtgtgtgatatgagcagttga	989
Db	1383	gtcaacatgtctgacggtcacccctgagtgctgatactatcgtgtgtgtgattgtgtgtct	1442
QY	990	gccccagtggt 1000	
Db	1443	gccaaagtgtgt 1453	
RESULT 8			
ABLO9105			
ID	ABLO9105	standard; cDNA; 1850 BP.	
XX	ABLO9105;		
AC	26-MAR-2002 (first entry)		
XX	Drosophila melanogaster expressed polynucleotide SEQ ID NO 21797.		
DT	Drosophila: developmental biology; cell signalling; insecticide;		
XX	Pharmaceutical; gene; ss.		
KW	Drosophila melanogaster.		
XX	WO200171042-A2.		
XX	27-SEP-2001.		
PD	23-MAR-2001; 2001WO-US09231.		
XX	23-MAR-2000; 2000US-191637P.		
XX	11-JUL-2000; 2000US-0614150.		
PR	(PEKE) PE CORP NY.		
XX	Venter JC, Adams M, Li PMD, Myers EW;		
PI	WPI; 2001-656860/75.		
XX	P-PSDB; ABB65002.		
DR	New isolated nucleic acid detection reagent for detecting 1000 or more		
XX	genes from Drosophila and for elucidating cell signalling and cell-cell		
PT	interactions -		
XX	Claim 1; SEQ ID NO 21797; 21np + Sequence Listing; English.		
PS	The invention relates to an isolated nucleic acid detection reagent		
CC	capable of detecting 1000 or more genes from Drosophila. The invention is		
CC	useful in developmental biology and in elucidating cell signalling and		

CC	cell-cell interactions in higher eukaryotes for the development of
CC	Insecticides,therapeutics and pharmaceutical drugs. The invention
CC	discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC	sequences (ABL01840-ABL16175) and the encoded proteins
CC	(ABB57737-ABB72072).
CC	The sequence data for this patent did not form part of the printed
CC	specification, but was obtained in electronic format directly from WIPO
CC	at ftp.wipo.int/pub/published_pcl_sequences.
XX	
SQ	Sequence 1850 BP; 430 A; 542 C; 485 G; 393 T; 0 other;
	Query Match 21.0%; Score 209.8; DB 23; Length 1850;
	Best Local Similarity 54.3%; Pred. No. 4.7e-50;
	Matches 527; Conservative 0; Mismatches 417; Indels 27; Gaps 4;
OY	36 gtaacagatttaaaccacaagtgtccaccacctaccaccccccggtgycgctttcct 95
DB	852 gcacgaagctccacgcctcagcacccggcagcagaagacagctccgccaccacacaccc 911
OY	96 ccaactcccacgctttagtcctccctacaccttgacacctgtccagctactctctgtga 155
DB	912 gccgaagctccacgacgcgcgcctccctccacacagctccgcgcctgcagctgcagca 971
OY	156 ccaaaagggaagggtgttgtttaagccctctcttgcaaaagaattgycagtagaagaaggat 215
DB	972 gttaagggcgtgtatatgtccatgtcccatgtgcgaagcgactagccgagugacaaaagctg 1031
OY	216 gatctatacaagaatgaagggaacgaagcagatgtgtagaatataccaagaagaatctgac 275
DB	1032 cgltc---acaaggcaaggcgagtggaatccatgtccatcataatctggtatcttga 1088
OY	276 tcttttgycctafgaagaagttgctcgtctcctccgagacgtgtgtgtgctctcccaagytct 335
DB	1089 ggccgaagaagcagcagccaaagcagctgtgcyggaacctgcaagcccccgagcagct 1148
OY	336 ggaatggcacccagttccctaacaggtgtcttccaaagatattccaatgaacaatttgtcg 395
DB	1149 ggaagcgcg-----ctacagagacattccggtgaccaaatgcgcgcga 1190
OY	396 gttatgtcacagcagattaatgatcatcaaaagacaacacatccattatattcttatc 455
DB	1191 ggtattgtccaaacgctcgtctgtgagtcacaagaacaactctccatctaactgaaccygt 1250
OY	456 gatgtaaataatavggagaagtttgtgtgtgacggaaagaactaaatagaatatgaagg 515
DB	1251 caatgccaaagtgtgaatacgtctgtgaagtccgcgcgcgaagggtgaacaagaagtacgaagg 1310
OY	516 agaaq---caaaatttcgtcaatfactcatcatataaaagcttcacgcttgcattgta 572
DB	1311 caggtgtgcccggtctctctgttaaaagactcatattacaaggccgtgcatctgcagctct 1370
OY	573 aaagttcccgaaagcaaatattctcttgatgtgacacaggtataagacaaatcatgttgt 632
DB	1371 aaggtctcccgaaagcgaactccgcctgtgatagtgacacagtaattcgcaagtaacgaagcgtc 1430
OY	633 gatgtcagtggtgcggtcaactctcctcctgaaggatcatatacaactatattgtttta 692
DB	1431 gatgtcttcggtgtgcgtctccacacagaacaggctctgtatacccgagctgtttccaatgvc 1490
OY	693 catataaaaggagtgaaaccattctctaattgattgtttcttlliaacaaacaaagcaaga 752
DB	1491 gcccgcaaggtgtctcctgtgaatctcttaagatgttcaaggcgtgtgcagcaaggcgcgc 1550
OY	753 gaggggtaaactcagccacatgaattccaggggtgtgcactttaacgtatcccaattatga 812
DB	1551 gacaaacaacttcacgcccacacgaattccagggtgtgacacatctcggttcttaattcgtgc 1610
OY	813 atgtgtggaattaaqaattctcctgtatattaaaccaacccaatgaacatgattttgga 872
DB	1611 atgttcogtgtgtgaacacagtttcgcgtgtcatcaaacccctccataatcgtgcatactctgc 1670
OY	873 attggtgtcttcagagataaactgttccctt---cagataatgaaaaagggtttgatgtg 929


```
OY 526 ttctgtcaatgactcatcataaagcttcaagcttggcagttgtaaaagttccgaag 585
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2495 tctctgtaaacgactcatcataaagcgtgcaagcttgcagcttaagttccgaag 2554
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 586 caaatctcttcttgatgagacagttataaagacaatcatgttcttgatgagcttg 645
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2555 cgaactccgcctgagatgagacagtaattcgcagagtaacgacgacgtgttcgttg 2614
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 646 cagtcagtaactcctcagacatcatcaacattgtgtttaaagacataaagag 705
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2615 ctgtctccacagacaagggctgattaccgcgattgtttcattgcgcgaaggggtg 2674
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 706 tggaaaccattgctaatgattgtgttctttagcaaccagaagcagaagggtaactac 765
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2675 tctctgagatctcttaaggaatgcaagcgtgcaagcgaagcgcgcgaacaacttc 2734
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 766 agccacatgaattccaggggtggaacttttagcatctccaattaggaatgttgaatta 825
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2735 agccccaagaaattccaggggtgcaacatctcgtttctaactcgtggaatgttcgtatg 2794
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
UY 826 agaat 830
    |||
DB 2795 gaat 2799
    |||

RESULT 11
ABL09104
ID ABL09104 standard; cDNA; 4983 BP.
XX
AC ABL09104;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 21794.
XX
KM Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PE 23-MAR-2001; 2001WO-US09231.
XX
PF 23-MAR-2000; 2000US-191637P.
XX
PY 11-JUL-2000; 2000US-0614150.
XX
AX (PEKE ) PE CORP NY.
PA
XX
PI Venter JC, Adams M, LI PMD, Myers EW;
XX
DR WPI; 2001-656860/75.
XX
DR P-PSDB; ABB65001.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Claim 1; SEQ ID NO 21794; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
```

```
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 4983 BP; 1461 A; 1129 C; 1029 G; 1364 T; 0 other:

Query Match 12.6%; Score 125.8; DB 23; Length 4983;
Best Local Similarity 63.3%; Pred. No. 1,5e-25;
Matches 193; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

OY 526 ttctgtcaatgactcatcataaagcttcaagcttggcagttgtaaaagttccgaag 585
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3333 tctctgtaaacgactcatcataaagcgtgcaagcttgcagcttaagttccgaag 3392
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 586 caaatctcttcttgatgagacagttataaagacaatcatgttcttgatgagcttg 645
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3393 cgaactccgcctgagatgagacagtaattcgcagagtaacgacgacgtgttcgttg 3452
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 646 cagtcagtaactcctcagacatcatcaacattgtgtttaaagacataaagag 705
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3453 ctgtctccacagacaagggctgattaccgcgattgtttcattgcgcgaaggggtg 3512
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 706 tggaaaccattgctaatgattgtgttctttagcaaccagaagcagaagggtaactac 765
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3513 tctctgagatctcttaaggaatgcaagcgtgcaagcgaagcgcgcgaacaacttc 3572
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
OY 766 agccacatgaattccaggggtggaacttttagcatctccaattaggaatgttgaatta 825
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3573 agccccaagaaattccaggggtgcaacatctcgtttctaactcgtggaatgttcgtatg 3632
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
UY 826 agaat 830
    |||
DB 3633 gaat 3637
    |||

RESULT 12
AAC46977
ID AAC46977 standard; DNA; 2348 BP.
XX
AC AAC46977;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 52106.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW Protein Identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PF 25-FEB-1999; 99US-0121825.
XX
PR 05-MAR-1999; 99US-0123180.
XX
PR 09-MAR-1999; 99US-0123548.
XX
PR 23-MAR-1999; 99US-0125788.
XX
PR 25-MAR-1999; 99US-0126264.
XX
PR 29-MAR-1999; 99US-0126785.
XX
PR 01-APR-1999; 99US-0127462.
XX
PR 06-APR-1999; 99US-0128234.
XX
PR 08-APR-1999; 99US-0128714.
XX
PR 16-APR-1999; 99US-0129845.
XX
PR 19-APR-1999; 99US-0130077.
XX
PR 21-APR-1999; 99US-0130449.
XX
PR 23-APR-1999; 99US-0130510.
XX
PR 23-APR-1999; 99US-0130891.
XX
PR 28-APR-1999; 99US-0131449.
XX
PR 30-APR-1999; 99US-0132048.
XX
PR 30-APR-1999; 99US-0132407.
```


PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-01324863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142920.
PR 13-JUL-1999; 99US-0142977.
PR 14-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.

PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151348.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 20-SEP-1999; 99US-0154039.
PR 22-SEP-1999; 99US-0154779.
PR 23-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.

XX AAA81500;
 AC
 XX 04-DEC-2000 (first entry)
 DT
 XX
 DE N. meningitidis partial DNA sequence gnm_47 SEQ ID NO:47.
 XX
 KW Neisseria meningitidis; Neisseria gonorrhoeae; genome; Immunogenic;
 KW antigen; vaccine; diagnosis; infection; antibacterial; identification;
 KW Meningococcus B; MenB; ds.
 XX
 OS Neisseria meningitidis.
 XX
 FN MO200022430-A2.
 PN
 XX 20-APR-2000.
 PD
 XX 08-OCT-1999; 99MO-US23573.
 PE
 XX 09-OCT-1998; 98US-0103794.
 XX 30-APR-1999; 99US-0132068.
 PR
 XX
 XX (CHIR) CHIRON CORP.
 XX
 XX Frazer CM, Hickey E, Peterson J, Tettein H, Venter JC;
 PI Maignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
 PI Rappuoli R, Pizza M;
 XX
 XX WPI; 2000-318079/27.
 DR
 XX
 PT Isolated nucleotide sequences of Neisseria meningitidis which can be
 PT used in the diagnosis and treatment of N. meningitidis infection and
 PT other Neisserial infections, for example, N.gonorrhoea -
 PT
 PS
 XX
 XX Claim 7; Page 1310-1317; 1760bp; English.
 CC
 XX
 CC The present invention describes methods of obtaining immunogenic
 CC proteins from Neisseria genomic sequences. AAA81453 to AAA82414
 CC represent specifically claimed Neisseria meningitidis genomic DNA
 CC sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
 CC Neisseria DNA sequences and their corresponding primers; AAA81254 to
 CC AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
 CC isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
 CC AAA81452 represent Neisseria meningitidis MenB polynucleotide ORF
 CC sequences, which are all used in the exemplification of the present
 CC invention. The nucleic acid sequences, protein sequences, and antibodies
 CC against them, can be used in the manufacture of a composition. The
 CC composition can be used as a medicament (or in the manufacture of a
 CC medicament) for treating, preventing or diagnosing infection due to
 CC Neisserial bacteria. For example, some of the identified proteins could
 CC be components of vaccines against Meningococcus B; against all serotypes;
 CC and/or against all pathogenic Neisseriae. Identification of sequences
 CC from the bacterium will also facilitate production of biological probes,
 CC particularly organism-specific probes. Attempts to make efficacious
 CC Meningococcus B vaccines have failed mainly due to antigen tolerance.
 CC Multivalent vaccines have also been tried but none have successfully
 CC overcome antigenic variability. The provision of further, complete
 CC sequences may provide an opportunity to identify secreted or surface
 CC exposed proteins that may be presumed targets for the immune system and
 CC which are not antigenically variable or at least more conserved than
 CC other more variable regions.
 CC
 XX Sequence 22863 BP; 5657 A; 5469 C; 5921 G; 5816 T; 0 other;
 XX

Query Match	8.88;	Score 88.2;	DB 21;	Length 22863;
Best Local Similarity	46.68;	Pred. No. 2.4e-14;		
Matches 420; Conservative	0;	Mismatches 448;	Indels 33;	Gaps 3

D8
QY 85 ccgcgttcctcacaatcccccagccttagctctcaacattcagcaccttgcacgcta 144
+ ++++++ | | | | | | | | | | | | | | | |
Db 8216 cagctgtactgcgcgtctgtaagaagccccagccgcgtctctctcagaagtgcgcccaagt 8275

[illegible]

DE	<i>Neisseria meningitidis</i> B nucleotide sequence SEQ ID NO:110.
XX	
KW	<i>Neisseria meningitidis</i> ; <i>Neisseria gonorrhoeae</i> ; immunogenic; vaccine; diagnosis; antigen; detection; infection; gene therapy; antibacterial;.

KM ds.
 XX Neisseria meningitidis.
 OS WO200066791-A1.
 PN 09-NOV-2000.
 PD 08-MAR-2000; 2000MO-US05928.
 PE 30-APR-1999; 99US-0132068.
 PR 08-OCT-1999; 99MO-US23573.
 PR 28-FEB-2000; 2000GB-0004695.
 XX (CHIR) CHIRON CORP.
 PA (GENO-) INST GENOMIC RES.
 PI Pizza M, Hickey E, Peterson J, Tettelin H, Venter JC, Masignani V,
 PI Galeotti C, Mora M, Ratti G, Scarcellini M, Scarlato V, Rappuoli R,
 PI Frazer CM, Grandi G;
 WP1: 2000-647603/62.
 XX Neisseria meningitidis B full length genome sequence and open reading
 PT frames are used to detect, treat and prevent Neisserial infections -
 XX
 PS Claim 7: Appendix A: 692pp: English.
 XX
 CC The present invention describes the full length genome of
 CC Neisseria meningitidis B (NMB). The sequences in AAF21544 and AAF21607
 CC to AAF21613 represent fragments of the NMB genomic sequence, as the
 CC sequence was too long to go in a record on its own it was split into 8
 CC sequences which overlap each other at the beginning and end of each
 CC sequence by 49980 bp (i.e. the last 49980 bp of AAF21544 is repeated at
 CC the beginning of AAF21607, the last 49980 bp of AAF21607 are repeated at
 CC the beginning of AAF21608, and so on). AAF21545 to AAF21588 encode the
 CC Neisseria proteins given in AAB58550 to AAB58593, and AAF21589 to
 CC AAF21606 represent PCR primers which are used in the exemplification of
 CC the present invention. The NMB genome and fragments from it have
 CC antibacterial activity, and can be used in vaccines and gene therapy.
 CC Neisseria nucleic acids, proteins and/or antibodies which binds to the
 CC proteins can be used in compositions for treating or preventing infection
 CC due to Neisserial bacteria or as a diagnostic reagent for detecting the
 CC presence of Neisserial bacteria or of antibodies raised to Neisserial
 CC bacteria. Computers, computer memory, computer storage medium or computer
 CC databases can be used in a search to identify open reading frames (ORFs)
 CC or coding sequences within the NMB genome. The DNA sequences provide
 CC further opportunities to find antigenic or immunogenic proteins which are
 CC more effective in vaccines than the outer membrane proteins currently
 CC used.
 XX
 SQ Sequence 349980 BP: 81351 A: 86755 C: 95584 G: 86290 T: 0 other:
 XX
 Query Match 8.8%; Score 88.2; DB 21; Length 349980;
 Best Local Similarity 46.6%; Pred. No. 1e-13;
 Matches 420; Conservative 0; Mismatches 448; Indels 33; Gaps 3;
 QY 85 cgcgtgtctctcaactccccagccttaagctcttaacttagccacccctgcagacta 144
 DB 70787 cagctctactatgcgcgtgaagcccgccgcgcgtctctgcagagctgcgccagctg 70846
 QY 145 ctctctgcgcaacaagggaggtgtllttagccctcttgcacaagaagtggcaagtag 204
 DB 70847 cgcctctgtctgtacaacaacaacgcgtatgtctgtgcgcgaactgtgtcgcg 70906
 QY 205 agaaaggatgtatctacaagtaaaaggagcaaggaacagatggtagaatcaacaaga 264
 DB 70907 agaccggtgttgacgtgaacgaltgcgcgtgcgcgtgtcgtatitgaag 70966
 QY 265 agatatgcactcttltgtcctagtaagttctcctgtccgagcgtgtgtgctc 324
 DB 70967 aagacgtacaagaatgcgcgtgcgaactcgcgagcgcgtgcctcgtgtgcacttc 71026

QY 325 ccacagctccttgaaatgagcaccagttcctacaggtgtcttcacagatcccaatcaga 384
 DB 71027 ctgcgcgcgcagcgtccttgaaagacgcg-----taccatgagcc 71065
 QY 385 acattcgtcgggttatgtacagcgcgttaataatcaacaagaacacatctatatt 444
 DB 71066 gctcgtgtcccggtgttgagaacgcctcttgcctctcaacaagaagaacgcatttga 71125
 QY 445 acccttctatgatgtaaatatggagaagtttgtgtgacggaagaactaataaga 504
 DB 71126 ctacattcaagaatgacaacaagaacacatcatgtgactgtcgttgagaagtacaaga 71185
 QY 505 taatt-----agaaggagaagcaaaattctgtcaatgacttcaatcaaaagcttcag 558
 DB 71186 aattcgagaagaacacacgcgcgtgaaactgtgcttattgtcttcttcgttaagaacgcgtg 71245
 QY 559 cttgtgcattttaaagtcccgaaagcaaatctctctgtgagatgagacagttataagc 618
 DB 71246 tgcgcgcctgaaanaatacccggtgtgtgaatgtctctgttgacggaagacatcgtgt 71305
 QY 619 aaatcatgtgtgtatgtcagttgtcagttgtcagtaactcctgcaggaactcaactca 678
 DB 71306 accacgctacttcgacatcgcgtatgcgaattgtgcagccacgcggttgtgtgccaa 71365
 QY 679 ttgtgttaatgacacataaaaggagtgaacccattgctaagtgttttcttag 738
 DB 71366 ttctgtgtatgtcgcgaacaaatgagcatltgcgcacatcgaaacgaatgtgtatcag 71425
 QY 739 caaccaaagcaagaagggttaaaactacagccatgaatccagggctggaactttaga 798
 DB 71426 cgaanaaagcaagaagcgcgaanaatcgcgtatcgagaatcgcagcgcggtgaacttaga 71485
 QY 799 tctcaatttaggaatgttgaatgaatgaattctctgtatataaccaccccaag 858
 DB 71486 ttaccacggcgggtacttctgtctatgtactaccccgatcaaacaccacccat 71545
 QY 859 catgtatttggcaattgtgtcctcagagataaactgtgcctgtcagataatgaaag 918
 DB 71546 ctgcgatttgggtatgacgcac-----taagagcgcgtgtgtgtgtaaaagcgc 71599
 QY 919 gttttagtggctatgacatgtctgttacactcagttgtgataccgggtgtgtgag 978
 DB 71600 aagttgtgtcgcgtcagatgatgtatctgtcctgtcttagcacaacgatatcatgag 71659
 QY 979 g 979
 DB 71660 g 71660

Search completed: August 13, 2002, 19:38:36
 Job time: 4333 sec

QY	738	gcacccaagcagagagagggtaactacagccacatggaattccagggtgggaactttag	797
Db	1193	gtggggaagcgttaagcaagcaacttccacccatgataaactctggaacttctact	1252
QY	798	atccccaatttaagaaagtttggaaatlaagaattctctgtctatattaaaccacctcaa	857
Db	1253	tctcgcgactcgcgtatggtttagtgatgaattgtagcgtattcttcgcgcgaagacaag	1312
QY	858	gcacgtattttgggaattggttgcttcaagagataacgtgtccctcgagataatgaataaa	917
Db	1313	gggtcattatagcgtctgttggagcgtcaagaagccaactgttgt---tgcgtataagagatbga	1369
QY	918	gggtttgatgtgctacgacatgacatgctgttaacactcaagtttgatcacccgggttgagat	977
Db	1370	tcttcctagtgtaaaaaacacacatgctgtggaatgtaactgcagatcatcgtcatltgtgatat	1429
QY	978	ggagcag 984	
Db	1430	ggagcgt 1436	

US-09-066-046-23
Sequence 23, Application US/09066046A
Patent No. 6204252

GENERAL INFORMATION:
APPLICANT: MURPHY, Cheryl

STOKER, James
BELTZ, Gerald A.

TITLE OF INVENTION: CHARACTERIZATION OF GRANULOCYTIC

NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:

ADDRESSEE: HALE AND DORR L.L.C.
STREET: 60 State Street

CITY: Boston
STATE: MA; COUNTRY: United States
; ZIP: 02109
;

COMPIER READABLE FORM
MEDIUM TYPE: FLOPPY
CONVERTED: TBM DO 2

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #10

```

; CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/066,046A
:

```

FILING DATE: 24-Apr-1998
 CLASSIFICATION: <Unknown>

NAME: Superko, Colleen

REFERENCE/DOCKET NUMBER: 106.941.1555
TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 526-6000
TELEFAX: (617) 526-5000

```

; INFORMATION FOR SEQ ID NO: 23:
;
; SEQUENCE CHARACTERISTICS:

```

```

;
; LENGTH: 2/06 base pairs
; TYPE: nucleic acid

```

TOPOLOGY: linear

HYPOTHETICAL: NO
ANTI-SENSE: NO

SEQUENCE	DESCRIPTION:	SEQ ID NO:
23		

US-09-066-046-23

Query Match	7.08;	Score 69.8;	DB 4;	Length 2706;
Best Local Similarity	46.88;	Pred. No. 6e-11;		
Matches 258; Conservative	0;	Mismatches 287;	Indels 6;	Gaps 1

Oy	378	atcagcaacattcgtcgggattgtgaacagcattatgaatcaataaacaacacatcc	437
Db	855	ATGAGCAAAATCCGCCAAGTTATAGCTGCTAGGCTTTAAGAGATCACAATAATCTCTGCT	914
Oy	438	cathattacccttcctalcsgatcgtataaalatgygaagaatttggtagcggaaagaaact	497
Db	915	ACACTCAGCAACCTTTATGTGAAAGTTGATAGACAAAGATGATGAGACTCAGAGCTAAGTAC	974
Oy	498	aatagaatatagaagggag-----aagcaaaattctgtcaatgaacttatcataaa	551
Db	975	AAAGATCCCTTTGGAAAGAGATGATGTTAAAGCTGGGGTTATGTCCTTTATACGA	1034
Oy	552	gcttcagcatttggcagatgttaaaagttcccgaaagcaaatcttccttgatgacacagtt	611
Db	1035	GCGGTGTGCTAGTACCTTTCCGAAATTTCCCTGCTGTAATGCGGAAATTTCAAGCGGATGAT	1094
Oy	612	ataagacaaatacatgltgttgatgctcagtgltgtcggatcactccgcagaatcacc	671
Db	1095	ATAGCTACAGGGACATTTGTAAACATTTGAGTGCGGGTAGTACCGAATAAGGGTTAGTG	1154
Oy	672	acacctattgtttaatatgcacataaagaaggtggaacacattgctcaatgagtgttc	731
Db	1155	GTGGCTGTTATCAGAMAGCGGGAACATATGCTCACTTGCGAATGGAACAGCAAGCACTGTT	1214
Oy	732	tctttgaacacaaagaagaagaggtgaactacagccacatgaaatccagtgagacat	791
Db	1215	GACTTAAAGTACAAAAGCAGAGAGTGGCAAGCTCTCTGTTTGATATGTGTGTGCACAC	1274
Oy	792	ttacagatclccaaattttaggaatglttgyaaatlaagaattctctcgtatattaaacca	851
Db	1275	TTTACTATTACAAATGATGTGTATGTATGAGTGCCTATGTTCTACCCCTTAATATCAACCT	1334
Oy	852	cctcaagcatgtatttggcacaattggtgtctcagaagataaactggtccctgcagatcat	911
Db	1335	CCTCAATGTGAACCTTGTGGTATGCATGCTATACAGCACAGCTCCTGTGGCAGATGATGT	1394
Oy	912	gaaaagagtt 922	
Db	1395	AAAGTAGAGAT 1405	

RESULT 3
US-08-975-762-41/c
; Sequence 41, Application US/08975762

GENERAL INFORMATION:

APPLICANT: Lodes, Michael J.
APPLICANT: Houghton, Raymond

```

; ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS AND
; ; NUMBER OF SEQUENCES: 73

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP

CITY: Seattle

COUNTRY: USA
ZIP: 98104

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY

```

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
;

```

```

; SOFTWARE:  PatecIII release #1.0,
; CURRENT APPLICATION DATA:
;

```

FILING DATE: 21-MAR-1984

ATTORNEY/AGENT INFORMATION
NAME: Mark David J.

REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.439

; REFERENCE/DOC#1 NUMBER: 21
; TELECOMMUNICATION INFORMATION:

TELEPHONE: 206-622-4900
 TELEFAX: 206-682-6031
 INFORMATION FOR SEQ ID NO: 41:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3073 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 US-08-975-762-41

Query Match 7.0%; Score 69.8; DB 4; Length 3073;
 Best Local Similarity 46.8%; Pred. No. 6.5e-11;
 Matches 258; Conservative 0; Mismatches 287; Indels 6; Gaps 1;

```

OY 378 atcagcaacattcgcggttattgacagcgatttaagcaataaagaaccataact 437
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  ^ 1210 ATGACCAAAATCCGCAAGTATAGCTGCTAGGCTTAAGAGCTCAAAATACCTCTGCT 1151
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    438 cattatccttctcatcgatgtaataatggaagaatttggtagcgaagaact 497
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  Db 1150 ACACACGACACCTTTAAAGAACTGATATGACCAAGTAGAGCTCAGAGCTAAGTAC 1091
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  OY 498 aataagataatgaaggag-----aagcaaaattctctgaatgacttatacaaa 551
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  Db 1090 AAAGATGCCCTTGTGAAAGGATGATGTTAAGCTTGGCTTATGCTCTTTATCAGA 1031
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  OY 552 gcttcagcttggcatglttaaaagttccgaagaacaattcttctgtagcagaagtt 611
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  Db 1030 GCGGTGTGCTAGTCTCTTCCGAAATTCCTGTGCTGAATGCGAGATTTCAGGCGATGAT 971
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  OY 612 ataagacaataatcattgltgtagtgcagtgltgcagtaacttctgcaggaactc 671
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  Db 970 ATAGCTACAGGACATATGTAAACATTGGAGTCGCGGTGATGACCATTAAGGGTAAAGT 911
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  OY 672 aacctatgtgtttaaigcacataaagaagtggaacattgtaagtatgtgt 731
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  Db 910 GTGCGTTATCAGAAAGCGGAACCTATGCTACTGCTGAATGGAAGCAAGCACTTGT 851
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  OY 732 tctttagcaaccaagaagaggttaactacagccacatgaattcagaagtggcact 791
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  Db 850 GACTTAAGTACAAAGCAAGTAGTGCAAGCTCTCTGTTTCGATATGCTGTGCAAC 791
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  OY 792 ttaagatcccaattgaagtgttgaatgaatgaatctctgctattataacca 851
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  Db 790 TTTACTATATCCAAATGCTGTGTATGAGGTGCTATTGCTAATTAATCAACCT 731
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  OY 852 cctcaagcatgtatttggcaattggtgctcagaagtaaacgtgctccgcagaataat 911
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  Db 730 CCTCAATCTGGAACTCTGGGTATGCTATGCTATACAGACAGCTCTGTGGCAGTAAATGGT 671
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  OY 912 gaaaaagggtt 922
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  Db 670 AAGCTAGAGAT 660
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

RESULT 4
 US-09-295-028-41/c
 ; Sequence 41, Application US/09295028
 ; Patent No. 6277381
 ; GENERAL INFORMATION:
 ; APPLICANT: Reed, Steven G.
 ; APPLICANT: Lodes, Michael J.
 ; APPLICANT: Houghton, Raymond L.
 ; APPLICANT: McNeill, Patricia D.
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS
 ; FILE REFERENCE: 210121.439C4
 ; CURRENT APPLICATION NUMBER: US/09/295, 028
 ; NUMBER OF SEQ ID NOS: 85
 ; SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 41
 ; LENGTH: 3073
 ; TYPE: DNA
 ; ORGANISM: Ehrlichia sp.
 US-09-295-028-41

Query Match 7.0%; Score 69.8; DB 4; Length 3073;
 Best Local Similarity 46.8%; Pred. No. 6.5e-11;
 Matches 258; Conservative 0; Mismatches 287; Indels 6; Gaps 1;

```

OY 378 atcagcaacattcgcggttattgacagcgatttaagcaataaagaaccataact 437
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  Db 1210 ATGACCAAAATCCGCAAGTATAGCTGCTAGGCTTAAGAGCTCAAAATACCTCTGCT 1151
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  OY 438 cattatccttctcatcgatgtaataatggaagaatttggtagcgaagaact 497
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  Db 1150 ACACACGACACCTTTAAAGAACTGATATGACCAAGTAGAGCTCAGAGCTAAGTAC 1091
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  OY 498 aataagataatgaaggag-----aagcaaaattctctgaatgacttatacaaa 551
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  Db 1090 AAAGATGCCCTTGTGAAAGGATGATGTTAAGCTTGGCTTATGCTCTTTATCAGA 1031
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  OY 552 gcttcagcttggcatglttaaaagttccgaagaacaattcttctgtagcagaagtt 611
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  Db 1030 GCGGTGTGCTAGTCTCTTCCGAAATTCCTGTGCTGAATGCGAGATTTCAGGCGATGAT 971
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  OY 612 ataagacaataatcattgltgtagtgcagtgltgcagtaacttctgcaggaactc 671
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  Db 970 ATAGCTACAGGACATATGTAAACATTGGAGTCGCGGTGATGACCATTAAGGGTAAAGT 911
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  OY 672 aacctatgtgtttaaigcacataaagaagtggaacattgtaagtatgtgt 731
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  Db 910 GTGCGTTATCAGAAAGCGGAACCTATGCTACTGCTGAATGGAAGCAAGCACTTGT 851
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  OY 732 tctttagcaaccaagaagaggttaactacagccacatgaattcagaagtggcact 791
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  Db 850 GACTTAAGTACAAAGCAAGTAGTGCAAGCTCTCTGTTTCGATATGCTGTGCAAC 791
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  OY 792 ttaagatcccaattgaagtgttgaatgaatgaatctctgctattataacca 851
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  Db 790 TTTACTATATCCAAATGCTGTGTATGAGGTGCTATTGCTAATTAATCAACCT 731
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  OY 852 cctcaagcatgtatttggcaattggtgctcagaagtaaacgtgctccgcagaataat 911
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  Db 730 CCTCAATCTGGAACTCTGGGTATGCTATGCTATACAGACAGCTCTGTGGCAGTAAATGGT 671
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  OY 912 gaaaaagggtt 922
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
  Db 670 AAGCTAGAGAT 660
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

RESULT 5
 US-09-106-582-41/c
 ; Sequence 41, Application US/09106582
 ; Patent No. 6306402
 ; GENERAL INFORMATION:
 ; APPLICANT: Reed, Steven G.
 ; APPLICANT: Lodes, Michael J.
 ; APPLICANT: Houghton, Raymond
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS AND
 ; NUMBER OF SEQUENCES: 73
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: SEED and BERRY LLP
 ; STREET: 6300 Columbia Center, 701 Fifth Avenue
 ; CITY: Seattle
 ; STATE: Washington
 ; COUNTRY: USA
 ; ZIP: 98104
 ; COMPUTER READABLE FORM:
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS

THERAP

QY 586 caaatctctcttgatgagacagatataagaacaaatcatgltgttgatgagtgctg 645
DB 3897 TTAGACCTCTTATCGACGGCGATGACGTGATTACCAACAATTTCCAGCTCAGCATGG 3956
QY 646 cggatgctctctgagacacatcacacccattggttgaatgacacataaaggag 705
DB 3957 CGGTTTCTACGGCCGGCCGGCTGGTAGCGCGGTCTGCGTGAATGCAACCTTGGA 4016
QY 706 tgaacacattgtaatgltgttctttagaaccaagaagaagggttaactac 765
DB 4017 TGGCAACATCGAGAGAAATCAAGAGACTGGCACTCAAGCGCTGACGGCACTGA 4076
QY 766 agccacatgaatccagggtgacatttaacatcccaatgaagaatggttgaaata 825
DB 4077 CCGTTGAATCTGACCGGGGTGATCACTCAACCAACGAGGTGTTGGTCC 4136
QY 826 agattctctgctatattaaaccacccaacacacatgatttggcaattggtctcag 885
DB 4137 TGAATCTACCGCGATCAACCCCGCAGAGCCCAATTCGGGTATCCACCGCTATCA 4196
QY 886 aggaataactgltccctgcagataaagaaagggttgaatggtgctgaatgctg 945
DB 4197 AAGATGTCGATGCGGTGAATGATCA-----GGTTGATCTGCGCATGATGTACC 4250
QY 946 ttacactcagttggtaccagggtggtgagtg 979
DB 4251 TGCGCTGTCTACGATCAACCGTCTGATGATGG 4284

RESULT 7
US-08-975-762-36
Sequence 36, Application US/08975762
Patent No. 6207169

GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Lodes, Michael J.
APPLICANT: Houghton, Raymond
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS AND
NUMBER OF SEQUENCES: 73
CORRESPONDENCE ADDRESS:
ADDRESS: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
City: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104

TREATMENT

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/975,762
FILING DATE: 21-MAR-1997
CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:
NAME: Makl, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.439
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-622-4900
TELEFAX: 206-682-6031

INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 1039 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Ehrlichia

US-08-975-762-36

Query Match 6.2%; Score 62.2; DB 4; Length 1039;
Best Local Similarity 47.8%; Pred. No. 6,4e-09;
Matches 181; Conservative 0; Mismatches 198; Indels 0; Gaps 0;

QY 544 tctaaagcttcagcttggtgacatgtttaaaggtcccgaaagcaaatctcttgatgag 603
DB 2 TAAAGACCGGTTGGTGGCTAGTCTTCCGAATTTCTTCTGTAATGCGGAGATTTTCAG 61
QY 604 acacagttataagacaacaaatcatgltgttgatgagtgctgagtgctgacag 663
DB 62 CGCATGATATGATCTACAGGAGCATATTGTAACATTGGAGTCCGGTAGTACCGATTAAG 121
QY 664 gactcaacacattggttgaatgacacataaaggagtgtaaacattgtctaag 723
DB 122 GGTAGTGTGCTGTATACAGAGCGGAAACTATGTCACTTGTGTGAATGGAGCAAG 181
QY 724 atgtgttctttagaacaagaagaagggttaactacagccacataatccag 783
DB 182 CACTTGTGACTTAAATACAAAAGCAAGAGCGCAAGCTCTGTGTTGATGATCTCG 241
QY 784 gtggcactttacgaltcccaattagaatgttggaaattgaagaattctctgtatla 843
DB 242 GTCAACCTTACTATTACCAATGCGTGTGATGGTCCGTAATTGCTACCCCTATTA 301
QY 844 ttaaccactcaagatgatttggcaattggtcttcagaggtaaacgtgctccg 903
DB 302 TCAACCTCTCTAATCTGGAATCTGGGTATGATGATCTATACAGCGCTCTGTGGCAG 361
QY 904 cagataagaaaagggtt 922
DB 362 TAGATGGTAAAGTAAAGAT 380

RESULT 8
US-08-821-324-36
Sequence 36, Application US/08821324
Patent No. 6231869

GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Lodes, Michael J.
APPLICANT: Houghton, Raymond
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS AND
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESS: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
City: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104

THERAP

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/821,324
FILING DATE: 21-MAR-1997
CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:
NAME: Makl, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.439
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-622-4900
TELEFAX: 206-682-6031

INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 1039 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

```

; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Ehrlichia
US-08-821-324-36

```

Query Match	6.2%	Score 62.2	DB 4	Length 1039
Best Local Similarity	47.8%	Pred. No. 6.4e+09		
Matches 181, Conservative	0	Mismatches 198	Indels 0	Gaps 0

Oy	tcataaagcttcagcgtttgcatglttaaaagttcccgaaagcaattcttcgtgattg	603
544	tcataaagcttcagcgtttgcatglttaaaagttcccgaaagcaattcttcgtgattg	
Db	2 TATACAGACGGGTGTGCTAGTCTCTTCCGAATCTCTGTGCTGAATGGCAATTTCAG	61
Oy	acacagcttataagacaaatcattgtttgtatgtcagtggttggtcaagtaaccttcag	663
604	acacagcttataagacaaatcattgtttgtatgtcagtggttggtcaagtaaccttcag	
Db	62 GCGATGATATGTCTACACGGGACATATGTATACCATTTGGAGTCCGGTAGTACCGATTACG	121
Oy	gactcatcacacattatglttlttaatgacatalaaagaggttgaaacacattgctaag	723
664	gactcatcacacattatglttlttaatgacatalaaagaggttgaaacacattgctaag	
Oy	122 GATTAGTGGTCCCTCTTTATTCAGAAAGACGGGAACATATGTCTCTTGTGGAATGGAGCAAG	181

724	atgtgtttcttcttagcaaccagaagagaggttaactacagccacatgaaatccagg	763
182	CACCTGTTTGACTTAAGTCAAAGCMAAGTGGCAAGCTCTGTTTGTGATATGTCTG	241
784	gtggacctttacgactcccaatttaggaatgtttggaattaagaatttcctgctcta	843
212	GTCACACCTTTACTATTTACCAATGCTGTGTGTATGGTCCCTATTGTCTACCCCTATTA	301
844	ttaacccacctcaagacatglatlittgacaaatggtgtcttcagagataaactggtccctg	903
302	TCAACCTTCCTCAATCTGGAACTCTTGCGTATGTCATACGACAGAGGCTCTGGCAG	361
904	cagataatgaaaaagagctt	922
362	TAGATGTGTAAGCTACGAT	380

```

1  RESULT 9
2  US-09-295-028-36
3  : Sequence 36, Application US/09295028
4  : Patent No. 6277381
5  : GENERAL INFORMATION:
6  : APPLICANT: Reed, Steven G.
7  : APPLICANT: Lodes, Michael J.
8  : APPLICANT: Houghton, Raymond D.
9  : APPLICANT: McNeill, Patricia D.
10 : TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS
11 : AND TREATMENT OF EIRLICHIA INFECTION
12 : FILE REFERENCE: 210121.439C4
13 : CURRENT APPLICATION NUMBER: US/09/295,028
14 : NUMBER OF SEQ ID NOS: 85
15 : CURRENT FILING DATE: 1999-04-20
16 : SOFTWARE: FastSeq for Windows Version 3.0
17 : SEQ ID NO: 36
18 : LENGTH: 1039
19 : TYPE: DNA
20 : ORGANISM: Ehrlichia sp.
21 : US-09-295-028-36

```

Query Match	6.28;	Score 62.2;	DB 4;	Length 1039;
Best Local Similarity	47.8%;	Pred. No. 6.4e-09;		
Matches 181; Conservative	0;	Mismatches 198;	Indels 0;	Gaps 0;

QY	544	tcataaagcctcagccttctgtgcagcttcaaaagctcccgagcaaatctctcttgatg	603
QY	2	tatcaacgacggtctgtcagtcctctccgaatccctgcgtcgaatgacggaattcag	61
Db	604	acacagcttaacgacaaatcatctgtttgtatgtcagtggttcggtcagtaactctcag	663
QY	62	gcgatgatatacttcacggagacatcttbaacatggagcttcggttcagtcgcataag	121

QY	664	gaccatcatcaaccacattggtgttttatatgcacataataaaagagtgagaaacattgtctaag	723
Db	122	ggttagtggtgccttgatcatcagaagacgagaaactatgtcacctgtcgtaaaatggagcaag	181
QY	724	atgtgttctctcttagcaaccacaagagagaggtaaacatacagccacatgaattccagg	783
Db	182	caattgttgacttaagtatacaaaagaagaagtgccaagctctctgttctcgatatgtctg	241
QY	784	gtgcacatttaacgatctccaaatttaggaatggtttggaattagaattctctcgcttta	843
Db	242	gtgcacaacttaccattaccatccaatgtgtgtgtatgtgttcgctatgtctacccctataa	301
QY	844	ttaacccaccccaagcatgtatatttttggaatttggtcttcacagagataaacctgttccctg	903
Db	302	tcaaccctcccatctcgaatctctgttgatgcatgtcatcagcagcgctccctgtggcag	361
QY	904	cagataatgaaaaaggtct	922
Db	362	tagatgtgataagtgtagaat	380

```

10 RESULT
11 US-09-106-582-36
12 : Sequence 36, Application US/09106582
13 : Patent No 6305402
14 :
15 : GENERAL INFORMATION:
16 :
17 : APPLICANT: Reed, Steven G.
18 : APPLICANT: Lodes, Michael J.
19 : APPLICANT: Houghton, Raymond
20 : TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DIAGNOSIS AND
21 : NUMBER OF SEQUENCES: 73
22 : CORRESPONDENCE ADDRESS:
23 : ADDRESSEE: SEED and BERRY LLP
24 : STREET: 6300 Columbia Center, 701 Fifth Avenue
25 : CITY: Seattle
26 : STATE: Washington
27 : COUNTRY: USA
28 :
29 : ZIP: 98104
30 :
31 : COMPUTER READABLE FORM:
32 : MEDIUM TYPE: Floppy disk
33 : COMPUTER: IBM PC compatible
34 : OPERATING SYSTEM: PC-DOS/MS-DOS
35 : SOFTWARE: PatentIn Release #1.0, Version #1.30
36 :
37 : CURRENT APPLICATION DATA:
38 : APPLICATION NUMBER: US/09/106, 582
39 : FILING DATE: 29-JUN-1998
40 :
41 : CLASSIFICATION:
42 :
43 : ATTORNEY/AGENT INFORMATION:
44 :
45 : NAME: Maki, David J.
46 : REGISTRATION NUMBER: 31,392
47 : REFERENCE/DOCKET NUMBER: 210121.439C2
48 :
49 : TELECOMMUNICATION INFORMATION:
50 :
51 : TELEPHONE: 206-622-4900
52 : TELEFAX: 206-682-6031
53 :
54 : INFORMATION FOR SEQ ID NO: 36:
55 :
56 : SEQUENCE CHARACTERISTICS:
57 : LENGTH: 1039 base pairs
58 : TYPE: nucleic acid
59 : STRANDEDNESS: single
60 : TOPOLOGY: linear
61 :
62 : MOLECULE TYPE: DNA (genomic)
63 :
64 : ORIGINAL SOURCE:
65 : ORGANISM: Ehrlichia
66 :
67 : US-09-106-582-36

```

Query Match	6.28;	Score 62.2;	DB 4;	Length 1039;
Best Local Similarity	47.88;	Pred. No. 6.4e-09;		
Matches 181; Conservative	0;	Mismatches 198;	Indels 0;	Gaps 0;

QY 544 tcataaagcttcagcttcttgcatgtttaaagttcccgaaagcaaatctcttgcgatcg 603

Db 2 TAATCAGAGCGGTTCGTAGTCCTTCGAAATTCGTGTCGTAATGCGAGATTTCAG 61
Oy 604 acacagatataagacaaatcatgltgtgtatgcatggtcgggtcagctcgtcag 663
Db 62 GCGATGATATAGCTTACAGGAGCTATGTACATTGAGATCGCGGTAGTACCGATTAAG 121
Oy 664 gactatcacacactatgtgttaatagcacataaagaagagtggaacactgttaaty 723
Db 122 GGTATGATGCTGCTGTATCAGAGAGCGGAATATGCACTCTCTGTAATGAGAGCAAG 181
Oy 724 atgtgttcttcttaagacccaagaagaaggttaacttaagacacatgattccagg 783
Db 182 CACTGTGACTTAACTAAGTACAAAAGCAGAGTGCAGAGCTCTCTGTTCTGATATGCTCG 241
Oy 784 gtgacattttagatctccaatlttagaatagtttggaattagaatttctctgtatla 843
Db 242 GTGCAACCTTACTATTAACCAATGATGATGATGATGATGATGATGATGATGATGATGAT 301
Oy 844 ttaccacacctcaagacatgatttttgccaattgtgtcctcagaagataaactgtccctg 903
Db 302 TCAACCTCTCTCAATCTGCAATCTTGCGTATGATGATGATGATGATGATGATGATGATGAT 361
Oy 904 cagataagaaaaggtt 922
Db 362 TAGATGCTAAGGTAGAT 380

RESULT 11
US-08-881-771A-5
Sequence 5, Application US/08881771A
Patent No. 6111071

GENERAL INFORMATION:
APPLICANT: Gershwin, Eric
APPLICANT: Leung, Patrick
APPLICANT: Coppel, Ross
TITLE OF INVENTION: RECOMBINANT FUSION PROTEIN
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:

COMPRISING PDC-

ADDRESSEE: Mandel & Adriano
STREET: 35 N. Arroyo Parkway, Suite 60
CITY: Pasadena
STATE: CA
COUNTRY: USA
ZIP: 91103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/881,771A
FILING DATE: 24-JUN-1996
CLASSIFICATION: 530
PRIOR APPLICATION NUMBER:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Adriano, Sarah B
REGISTRATION NUMBER: 34,470
REFERENCE/DOCKET NUMBER: 30448.51US01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 626-395-7801
TELEFAX: 626-395-0694
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1096 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO

FRAGMENT TYPE:
ORIGINAL SOURCE:
US-08-881-771A-5

Query Match 4.9%; Score 48.6; DB 3; Length 1096;
Best Local Similarity 85.7%; Pred. No. 7.8e-05;
Matches 54; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Oy 1 atatacagcatlctgactataagccacaggaagtaacagatttaaacacacagtcg 60
Db 755 ATATATACACATTTCCTGCTACTATAGCCACAGCAAGTAAACATTTAGCATTCCTCCGGGTC 814
Oy 61 cac 63
Db 815 GAC 817

RESULT 12
US-08-881-771A-8
Sequence 8, Application US/08881771A
Patent No. 6111071

GENERAL INFORMATION:
APPLICANT: Gershwin, Eric
APPLICANT: Leung, Patrick
APPLICANT: Coppel, Ross
TITLE OF INVENTION: RECOMBINANT FUSION PROTEIN
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:

COMPRISING PDC-

ADDRESSEE: Mandel & Adriano
STREET: 35 N. Arroyo Parkway, Suite 60
CITY: Pasadena
STATE: CA
COUNTRY: USA
ZIP: 91103
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/881,771A
FILING DATE: 24-JUN-1996
CLASSIFICATION: 530
PRIOR APPLICATION NUMBER:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Adriano, Sarah B
REGISTRATION NUMBER: 34,470
REFERENCE/DOCKET NUMBER: 30448.51US01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 626-395-7801
TELEFAX: 626-395-0694
TELEX:
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 426 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
US-08-881-771A-8

Query Match 4.8%; Score 48.4; DB 3; Length 426;
Best Local Similarity 98.0%; Pred. No. 5.1e-05;
Matches 49; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 atatacgcatttctgactatagccacccgaagtaacatttaaa 50
|||||
Db 374 AATATACAGATTCTGACTATAGCCACCGAGGTAACAGATTIAGCA 423

RESULT 13

US-08-209-747-1/c
; Sequence 1, Application US/08209747
; Patent No. 5733771
; GENERAL INFORMATION:
; APPLICANT: Lewis, Randolph V.
; APPLICANT: Colgin, Mark
; TITLE OF INVENTION: cDNAs Encoding Minor Ampullate Spider
; TITLE OF INVENTION: Silk Proteins
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-3487
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/209,747
; FILING DATE: 14-MAR-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 1447-104P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-205-8000
; TELEFAX: 703-205-8050
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2793 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Nephila clavipes
; TISSUE TYPE: minor ampullate gland
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 183..2675
; OTHER INFORMATION: /product= "N. clavipes minor
; OTHER INFORMATION: ampullate silk protein"
US-08-209-747-1

Query Match 4.3%; Score 42.6; DB 1; Length 2793;
Best Local Similarity 59.5%; Pred. No. 0.0085;
Matches 72; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

OY 38 aacagatttaaaccaagtgccacactaccaccccggtgagcgctgttcctcc 97
|||||
Db 1663 ACCATATCCAGCACACCGGTGCTGCACTGCCACGACGCGCACACAGCTCCAGCAC 1604
OY 98 aactcccaagccttagctctacaccttaagcaccctgcccagctactctgtgagcc 157
|||||
Db 1603 AGCAGCTTACCTTAACACACGCGCTCGGCTCTGTTGACGACGCTGCACAGCAGACC 1544
OY 158 a 158
Db 1543 A 1543

RESULT 14

US-08-458-298-1/c
; Sequence 1, Application US/08458298
; Patent No. 5756677
; GENERAL INFORMATION:
; APPLICANT: Lewis, Randolph V.
; APPLICANT: Colgin, Mark
; TITLE OF INVENTION: cDNAs Encoding Minor Ampullate Spider
; TITLE OF INVENTION: Silk Proteins
; NUMBER OF SEQUENCES: 56
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-3487
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/458,298
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/209,747
; FILING DATE: 14-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M.
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 1447-104P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-205-8000
; TELEFAX: 703-205-8050
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2793 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Nephila clavipes
; TISSUE TYPE: minor ampullate gland
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 183..2675
; OTHER INFORMATION: /product= "N. clavipes minor
; OTHER INFORMATION: ampullate silk protein"
US-08-458-298-1

Query Match 4.3%; Score 42.6; DB 1; Length 2793;
Best Local Similarity 59.5%; Pred. No. 0.0085;
Matches 72; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

OY 38 aacagatttaaaccaagtgccacactaccaccccggtgagcgctgttcctcc 97
|||||
Db 1663 ACCATATCCAGCACACCGGTGCTGCACTGCCACGACGCGCACACAGCTCCAGCAC 1604
OY 98 aactcccaagccttagctctacaccttaagcaccctgcccagctactctgtgagcc 157
|||||
Db 1603 AGCAGCTTACCTTAACACACGCGCTCGGCTCTGTTGACGACGCTGCACAGCAGACC 1544
OY 158 a 158
Db 1543 A 1543

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 13, 2002, 17:33:53 ; Search time 1695.31 Seconds
(without alignments)
7969.311 Million cell updates/sec

Title: US-09-761-580-1_COPY_800_1800

Perfect score: 1001
Sequence: 1 atatacagcattgctgac.....cagltggagccagtgacct 1001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Archived: 13736207 seqs, 6748477542 residues

al number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hnc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hnc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	739	73.8	751	AL040276	AL040276 DKF2P434K
2	712.2	71.1	745	AL581280	AL581280 AL581280
3	685.4	68.5	738	AU126868	AU126868 AU126868
4	682.2	68.2	833	AU127045	AU127045 AU127045
5	630.6	63.0	673	BG778185	BG778185 602666342
6	630	62.9	833	BI092658	BI092658 602858446
7	609.6	60.9	901	BI730459	BI730459 603350793
8	595	59.4	1030	BM463514	BM463514 AGCWCOURT
9	582.4	58.2	857	BG253201	BG253201 602364158
10	555.6	55.5	720	BF939962	BF939962 nacc5a04
11	552.8	55.2	680	BG503419	BG503419 602550846
12	532.4	53.2	1125	AA867881	AA867881 vx22c10.r
13	526	52.5	664	BG390485	BG390485 602416136
14	515.4	51.5	654	BB850747	BB850747 ux01e11.y
15	509	50.8	774	BI658188	BI658188 603285296
16	508	50.7	669	BE627518	BE627518 u052f03.y
17	473.2	47.3	828	BF209700	BF209700 601874236

18	451	45.1	790	10	BI836946	BI836946 603084267
19	441.8	44.1	574	10	BG087041	BG087041 H3133G11-
20	433.4	43.3	471	10	N99373	N99373 yz96a07.r1
21	428.4	42.8	708	9	AM950165	AM950165 EST362130
22	414.4	41.4	444	9	AU160458	AU160458 AU160458
23	407.6	40.7	909	10	BE784101	BE784101 601471302
24	405.2	40.5	911	10	BF796701	BF796701 602258322
25	403.6	40.3	601	9	AA445205	AA445205 v163d04.r
26	399.4	39.9	1056	10	BM455483	BM455483 AGCWCOURT
27	395.6	39.5	583	10	BF562783	BF562783 UT-R-BJOP
28	395	39.5	584	9	AI419467	AI419467 ttf16e01.x
29	394.2	39.4	575	9	AY667649	AY667649 AY667649
30	391	39.1	918	10	BF144153	BF144153 60178848
31	390.4	39.0	563	9	AA275794	AA275794 vc25h09.r
32	388.4	38.8	582	9	AJ393037	AJ393037 AJ393037
33	384	38.4	488	9	BE137467	BE137467 u053g08.y
34	383.4	38.3	809	10	BI733158	BI733158 603354594
35	378.8	37.8	463	10	BE629397	BE629397 u037e01.y
36	378.6	37.8	567	10	BM236071	BM236071 K0428G04-
37	362.4	36.2	463	10	N77305	N77305 yv43e08.r1
38	361.2	36.1	627	10	BI333861	BI333861 602993980
39	360.6	36.0	479	9	AA512816	AA512816 v013e02.r
40	360.2	36.0	519	9	AA288380	AA288380 m747b11.r
41	358.4	35.8	520	9	AA288386	AA288386 m747b11.r
42	355.6	35.5	580	9	AA543438	AA543438 v037g12.x
43	354	35.4	564	10	BF725774	BF725774 BX19g09.y
44	351	35.1	688	9	AL657557	AL657557 AL657557
45	348.4	34.8	352	9	AA315214	AA315214 EST187073

ALIGNMENTS

RESULT 1
AL040276 751 bp mRNA linear EST 29-FEB-2000
LOCUS DKF2P434K2113.r1 434 (synonym: hncs3) Homo sapiens CDNA clone
DEFINITION DKF2P434K2113 5', mRNA sequence.
ACCESSION AL040276
KEYWORDS EST.
SOURCE Homo sapiens
ORGANISM human.
REFERENCE Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
AUTHORS Koehner, K., Beyer, A., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
TITLE EST (Koehner, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: Koehner K
MIPS
Am Klopferspitz 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ), Email: s.wiemann@dkfz-heidelberg.de;
sequenced by BMFZ (Biomedical Research Center at the Charité,
Berlin/Germany) within the cDNA sequencing consortium of the German
Genome Project.
No si sequence available.
This clone (DKF2P434K2113) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

source 1..751
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKF2P434K2113"
/clone_1lb="434 (synonym: hncs3)"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pSPORT1; Site_1: NotI; Site_2: SalI"
BASE COUNT 226 a 147 c 218 t

QY	883	caaggagataaacctgtgtccctccgcagaataatgaaaagggttgcagtcgtacgtatgt	942
Db	145	CAGAGGATAACTGTGTCCCTCCACATTAATGAAAAAGGTTTGATCTGCTACGATGATGT	86
QY	943	ctgtaccctcagttgtgtatcacccggtgtgtgtgtagtcagtcagtcggaccagtgcctt	1001
Db	85	CTGTACACTCAGTGTGTATCMCCGGGTGTGTGATGTAGAGCATTTGAGCCACGAGGCTT	27
RESULT	3		
LOCUS	AU126868	738 bp	linear
DEFINITION	AU126868 NT2RP2 Homo sapiens CDNA clone NT2RP2000113 5', mRNA		EST 23-Oct-2000
ACCESSION	AU126868		
VERSION	AU126868.1		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
	Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y.,		
	Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and		
	Isogai, T.		
TITLE	HRI human CDNA project		
JOURNAL	Unpublished (2000)		
COMMENT	Contact: Takao Isogai		
	Genomics Laboratory		
	Helix Research Institute		
	1532-3 Yana, Kisarazu, Chiba 292-0812, Japan		
	Tel: 81-438-52-3951		
	Fax: 81-438-52-3952		
	Email: genomese@hri.co.jp		
	HRI human CDNA project; 5'- & 3'-end one pass sequencing; Helix		
	Research Institute; CDNA library construction; Department of		
	Virology, Institute of Medical Science, University of Tokyo, and		
	Helix Research Institute.		
FEATURES			
SOURCE	Location/Qualifiers		
	1..738		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="NT2RP2000113"		
	/clone_1lb="NT2RP2"		
	/cell_type="teratocarcinoma"		
	/cell_line="NT2"		
	/note="Vector: pME185FL3; mRNA from NT2 neuronal precursor		
	cells after 2-weeks retinoic acid (RA) induction"		
JE COUNT	226 a 171 c 141 g 192 t		8 others
UNIGIN			
Query Match	68.5%; Score 685.4; DB 9; Length 738;		
Best Local Similarity	97.0%; Pred. No. 3.8e-167;		
Matches	715; Conservative 0; Mismatches 19; Indels 3; Gaps 2;		
QY	21	tattagcgcaacggaagtaacagatataaacaacacaagtcgcacacactaccaccccg	80
Db	1	TATAGGCCAACGGAATTAACAGATTAAAAACCAAGCGCCACACACTACCCACCCCG	60
QY	81	gtggccgctgttctccaaactccccaagcctttagctctctacaccttaagcaccctgcca	140
Db	61	GGGGCGCGTGTTCCTCAACTCCCGCGCTTATAGCTCTACACCTTCACACCCCTGCCA	120
QY	141	gtctaccctcgtcggaaccaagggaaggggttgttagccctctgcaagaagtggca	200
Db	121	GCTACTCTCGCTGACCAAAAGGAGGGGTGTGTGTTGTTAGCCCTCTTGCAAAAGATTGGCA	180
QY	201	gttagaaaggagatattatcttacaacaagtaaaaggagacagacagatgtgaatcacc	260
Db	181	GTAAGAAAGGAGTTGATCTTTACACAAGTAAGGAGGACAGACAGATGTTAGAAATCACC	240
QY	261	aagaagatatcgactcttltgtgctagttaaagttgctcctgctccggcagctgttg	320

Db	241	AAGAAAGATATCGACTCTTTGTGCTAGTAAAGTGTGCTCTGCTCCGGCAGCTGTGTG	300
Qy	321	ccctcccaaggtccttggaatgagccacagttcctcaagtgcttcacagatatccaatc	380
Db	301	ctctcccaaggtccttggaatgagccacagttcctcaagtgcttcacagatatccaatc	360
Qy	381	agcaacatcgctcggttatctgcacacgataatgcaatcaatcaaacacatccat	440
Db	361	agcaacatcgctcggttatctgcacacgataatgcaatcaatcaaacacatccat	420
Qy	441	tattaccttctcatcgatgtaataatgaggagaagtttctgtgtgtaacggaagaaat	500
Db	421	tattaccttctctatcaatgtaataatgaggagaagtttctgtgtgtaacggaagaaat	480
Qy	501	aagatattagagaaggagaagaacaattctgtcaatgtaattcatcataaaagctcagct	560
Db	481	aagatattagagaaggagaagaacaattctgtcaatgtaattcatcataaaagctcagct	540
Qy	561	ttagcatgtttaaaagttcccgaaacaattctcttgatgtagacagttataagaca	620
Db	541	ttggcatgttttaaaagttcccgaaacaattctcttgatgtagacagttataagaca	600
Qy	621	aatactgttgtagtgcagtggttggtgagttacgtcctgcagagctatacagctatt	680
Db	601	aatactgttgtagtgcagtggttggtgagttacgtcctgcagagctatacagctatt	660
Qy	681	gtgttcaatgacat-ataaagagtggaacaacatgtctaag-attgtttcttca	737
Db	661	gtgttcaatgacat-ataaagagtggaacaacatgtctaag-attgtttcttca	720
Qy	738	gcaacccaagaagaga	754
Db	721	gcaacccaagaagaga	737

LOCUS AUI127045

DEFINITION AUI127045 NT2RNP2 Homo sapiens cDNA clone NT2RNP2000456 5', mRNA sequence.

ACCESSION AUI127045

VERSION AUI127045.1

KEYWORDS GI:10951761

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

AUTHORS Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and Isogai,T.

TITLE HRI human cDNA project

JOURNAL Unpublished (2000)

COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3952
Fax: 81-438-52-3952
Email: genom@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.

FEATURES

Source

1..833

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="NT2RNP2000456"

/clone_1ib="NT2RNP2"

/cell_type="teratocarcinoma"

/cell_line="NT2"

Location/Qualifiers

/note="Vector: pME18SFL3; mRNA from NT2 neuronal precursor cells after 2-weeks retinoic acid (RA) induction"

BASE COUNT 249 a 192 c 165 g 220 t 7 others

ORIGIN

Query Match 68.2%; Score 682.2; DB 9; Length 833;
Best Local Similarity 95.7%; Pred. No. 2.7e-166;
Matches 711; Conservative 0; Mismatches 30; Indels 2; Gaps 1;

21 tatagcaaccgaagtaacagattaaacacaaagtcacacccaccacccccc 80
1 TATAGCCCAACCGAAGTAACAGATTAAACCAAGGCCACCACTTACCCACCCCG 60
81 gtagccgctgttccctccacacccacccacccacccacccacccaccc 140
61 GTGGCCCTGTTCTCCCAACTCCCAAGCTTACCTCCTACACCTTACGACCCGCCA 120
141 gctactcctgtgacccaaaggaggtgttltgttagccctcttgcaagaagtgtgca 200
121 GCTACTCCTGCTGACCAAAAGGAGGGGTGTTGTTAGCCCTTGCAAAAGATTGGCA 180
201 gtaggaagagagattgattctacacaaaggaagacagacagatgtatgaatcacc 260
181 GTAGAGAAAGGATGATCTTACACAGTAAAGGACAGGACGATGTGAAATCACC 240
261 aagaagatatacgaactcttctgtcctagtaagttgtctcctgtccgagctgtgtg 320
241 AAGAAGATATCGACTCTTTGTGCTAGTAAAGTGTCTGCTCCGACCTGTGTG 300
321 cctccacaggtccctgggaatggcaccagttctacaggtgtcttccagaatcccaatc 380
301 CCTCCACAGGTCCTGGAAATGACACGATTCACAGGTCTTCCACAGATATCCCAATC 360
381 agcaacatcgtcgggttattgacagcagattaatgcaatcaagaacacatacctcat 440
361 AGCAACATTCGTCGGGTATTGTCAGACGATTAATCAATCAAGCAACATACCTCAT 420
441 tattacccttctatcgatgataataggaagaagtttctgttgaaggaagaactaat 500
421 TATTACCTTCTATCAATTAATATGAGAGAGATTTGTGTGTCAGAAAGCAATTAAT 480
501 aagatattgaagggaggaagcaaatcttctgcaatgacttcaataaagacatcagct 560
481 AACATATTGAAGGAGCAACAAATTTCTGTCAATGACTTCATTAAGGTTACGCT 540
561 ttggcattgttaaaagttcccggaagcaaatcttctgtgacaggtatataagaaa 620
541 TTGGCATGTTTAAAGTCCCGAAGCAAAATCTTGTGATGACACAGTTATGANAACA 600
621 aatcagttgttgatgacagttgtgtgacagtaactcctgacagatcacaactatt 680
601 AATCAATGTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
681 gtagttaaagacataa-aaagagtggaacacatgctaaatgagtgttctttag 738
661 GTGTTTAAAGCACATTAATAAAGGAATGGAACCAATGCTAATGAATGTTGTTCTT 720
739 caaccaaagagagagaggttaa 761
721 TANCACACCCCAAAAGCCAAAGAA 743

RESULT 5
LOCUS BG778185 673 bp mRNA linear EST 15-MAY-2001
DEFINITION 60266342F1 NIH_MGC_50 Homo sapiens cDNA clone IMAGE:4806083 5',
ACCESSION BG778185 mRNA sequence.
VERSION BG778185.1 GI:14048502
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Mammalia; Euthelia; Primates; Catarrhini; Homnidae; Homo.
TITLE NIH-MGC http://mgc.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs@emil.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNML at:
http://image.llnl.gov
Plate: LNCMI655 row: c column: 12
High quality sequence stop: 667.

FEATURES
source
1..673
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4806083"
/clone_1lb="NIH_MGC_60"
/tissue_type="adenocarcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: prostate; Vector: pDNR-LIB (Clontech); Site_1: SfiI (ggccattagggc); Site_2: SfiI (ggccattagggc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGGCGGCGCCGACATG-dt(30)BN-3' (where B = A, C, G or T). Average insert size 1.5 kb (range 0.9-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH-MGC Library."

BASE COUNT 193 a 168 c 139 g 173 t

Query Match 63.0%; Score 630.6; DB 10; Length 673;
Best Local Similarity 99.0%; Pred. No. 6.2e-153;
Matches 666; Conservative 0; Mismatches 4; Indels 3; Gaps 3;

28 caacggagtaacagattaaacacaaagtcacacacccacccacccggtggcg 87
1 CAACGGAAGTAACAGATTAAACCAAGCGCCACCTACACCCCGGTGGCG 60
88 ctgtctcccaactccacagctttagctctacacacttcaagacccctgccaagctactc 147
61 CGTTCTCTCAACTCCCAAGCTTTAGCTCTACACCTTACACACCTGCTCCAGCTACTC 120
148 ctgtggacacaaaggaaggtgtttagccctcttcaagaagaatttggagaaga 207
121 CTGCTGGACCAAGGAGAGGTTGTTGTTAGCCCTTCAAGAAAGTTGGAGTAGAGA 180
208 aagggattgattctacacaaaggaaggaagacagatgtgataatcacaagaag 267
181 AAGGATTTATCTTACAAAGTAAAGGACAGACCAATGTGTAATCACAAGAAAGG 240
268 atacgaactccttctgtcctagtaagttgtctcctgtcctcggcagctgttgcctcca 327
241 ATATCGACTCTTGTGCTTAAAGTGTCTCTGCTCCGACGCTGTTGCTCCCA 300
328 caggtccttgatggacaggttctcagaaggtgtcttcaagatcccaatcagaaca 387
301 CAGGCTCTGGAATGGACCAAGTTCTTACAGGTGTCTTACAGATATCCAAATCAGACACA 360
388 ttctgtcgggtattgacagcagattaatgcaatcaagaacacatattatc 447
361 TTCTGCTGGGTTATGACAGCGATTAAATGCAATCAAGCAACATTAATATAC 420

Oy 448 ttctcatcgtatgaataatgaggaagtttgctgtagtcggaagaacttaataagatat 507
|||||
Db 421 TTTCTATCATGTACATATGAGGAGAG--TTTGTGTGAGGAGAACTTAATAAGATAT 479
|||||
Oy 508 tagaaggagagaagaataatctctgcaatgactcaataaagcttagctggcat 567
|||||
Db 480 TAGAAGGAGGAGAGCAAAATTTCTGTCATGACCTCATCTAAAGCTTCAGCTTGGCAT 539
|||||
Oy 568 gtttaaaagttccggaagaactctcttgatgagacagttataaagaa-aatcat 626
|||||
Db 540 GTTTAAAGATTCCAGAGCAAAATTTCTTGTGATGACACACTTATAGACAAATCAT 599
|||||
Oy 627 gttgttga-tgctagtgctgcgtagtactctctgcaaggaactcaactatgtgt 685
|||||
Db 600 GTTGTCAGCTGACGTGTCGGTCTGCTCTGACAGGACTCATCACACTATTGTGT 659
|||||
Oy 686 taatgcacatata 698
|||||
660 TAATGCACATTTTA 672

RESULT 6
B1092658 833 bp mRNA linear EST 20-JUN-2001
LOCUS 60285844671 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:499811 5',
DEFINITION mRNA sequence.
ACCESSION B1092658
VERSION B1092658.1 GI:14510988
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLES NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov

Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM1030 row: k column: 12
High quality sequence stop: 746.
Location/Qualifiers

FEATURES
source 1..833
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:499811"
/clone_1lb="NIH_MGC_10"
/cell_line="MGC36"
/lab_host="DH10B"
/note="Organ: cervix; Vector: pCMV-SPORT6; Site: 1; NotI;
Site: 2; SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.5 kb. Library prepared by Life
Technologies."

BASE COUNT 252 a 162 c 246 t
ORIGIN

Query Match 62.9%; Score 630; DB 10; Length 833;
Best Local Similarity 98.9%; Pred. No. 9, 6e-153;
Matches 645; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

Oy 352 ctacaagtgctctcacagatcccaatcagacaatcgtcgggtatattgacagcgat 411
|||||
Db 1 CTACAGGTGCTCTCACAGATATCCCAATCAGCAACATGTCGGGTATTATGACAGGAT 60
|||||
Oy 412 taatgcaatcaagaacaaccatactcattatatacctctcagatgtaataatggag 471

|||||
Db 61 TAATCAATCAAGCAAAACCATACCTCATATTACTTCTCATGATTAATATGGGAG 120
|||||
Oy 472 aagtttggtaggaagaagacttaataagattagaaggaggaagcaaatctcg 531
|||||
Db 121 AAGTTTGGTGTAGGAGAAAGATTAAATTAAGATTATTAAGGAGAGAACCAAAATTTG 180
|||||
Oy 532 tcaatgacttcataaagaagcttagcttgatgtagttaaagtcccgagaanaat 591
|||||
Db 181 TCAATGACTTCAATTAAGAGCTTCAGCTTGGCATGTTAAAGTCCGAGCAATTT 240
|||||
Oy 592 ctcttgatgtagcaagattataagaacaatcagtgctgtagtgcagtggtgca 651
|||||
Db 241 CTTCCTGGATGACACAGTATTAAGCAAAATCATGCTTGTATGTCAGCTGGGCTCA 300
|||||
Oy 652 gtacttcgtaggaactcacaacctatgtgttaaatgacataaaggagtcgaaa 711
|||||
Db 301 GTACTCCTCAGAGACTCATACACCTATTGTGTTTAATCATATTAAGAGAGTGAGAA 360
|||||
Oy 712 ccattgctaagatgctgttctcttagcaaccaagaagagaggttaactacagcac 771
|||||
Db 361 CCATTCTAATGATGTTGTTCTTTAGCAACCAAGCAAGAGGAGTAATACAGCCAC 420
|||||
Oy 772 atgaattcagaggtggcacttttagatctcgaattagaagtgttgaattagaat 831
|||||
Db 421 ATGAATTCAGAGGTGACCTTTAGCATCTCCATTTAGGAATGTTGAATTAAGAAAT 480
|||||
Oy 832 tctctgtaattlaaaccaacctcaagcagtatttctgcaattggtgctcagaagata 891
|||||
Db 481 TCTCTCTATTATTAAACCACTCAAGCATGATTTTGGCAATTTGGTCTCAGAGAGATA 540
|||||
Oy 892 aactgtccctcagacata--atgaagaaggttgatgctgtagatgtagtctgttc 949
|||||
Db 541 AACTGTCCCTCGAGATTAAGAACACAGAGGTCGATGTCGATGATATGTCGTGTAC 600
|||||
Oy 950 actcaatgtgataccagcgggtgtagatggaagcagtggaagccagtgctt 1001
|||||
Db 601 ACTCACTTGTGATACACCGGCTGTGTGATGAGACAGTTGAGACCCAGCTGCTT 652

RESULT 7
B1730459 901 bp mRNA linear EST 20-SEP-2001
LOCUS 603350793F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:5358430 5',
DEFINITION mRNA sequence.
ACCESSION B1730459
VERSION B1730459.1 GI:15707472
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
TITLES NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov

Tissue Procurement: The Cepko Laboratory
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LHAM1911 row: 1 column: 23
High quality sequence stop: 843.
Location/Qualifiers

FEATURES
source 1..901
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:5358430"
/clone_1lb="NIH_MGC_94"

/tissue-type="retina"
/lab-host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dr primed.
Average insert size 3.3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH-MGC Library."
BASE COUNT 243 a 196 c 225 g 237 t
ORIGIN

Query Match 60.9%; Score 609.6; DB 10; Length 901;
Best Local Similarity 86.4%; Pred. No. 2e-147;
Matches 731; Conservative 0; Mismatches 109; Indels 6; Gaps 5;

```
OY 162 ggaagaggttctgttcccttcgcaagaagttgagagagagattgattcctt 221
DB 1 GGAAGAGGTTCGTGTTAGTCTCTTGGAAGAGTTGGACGAGAGAGATTGACCTC 60
OY 222 acacaagtaaaagagacagacagatgttagatcaccaagaagatatacactttt 281
DB 61 ACCCAAGTTAAAGGACGGGACCGACAGACAGATATCATACAGAGACATTGACTTTT 120
OY 282 gtgctctagaagttgtctctcctccgagctgtgtgtcctccacagctcctggaatc 341
DB 121 GTGCTCTTAAAGGCTCTCTCTGTCAGACGCTGCATGCTCCCGGGTCCAAAGATG 180
OY 342 gacacagctcccaagagttcttcaagatatacccaatagcaacattcgtcggtatc 401
DB 181 GCACACGCTCTCGACAGGTCTTCAACAGACATCCCATAGCAACATTCGTGAGAGAT 240
OY 402 gcaacagatlaaagaacaaagacacacatcattatcattcattcattcattcatt 461
DB 241 GCCCAAGAGCTCATGACAGTCGAAAGCAGACTATACCTATTATCTGTCATATGTA 300
OY 462 aatataagagagagtttltgttgaagaaagaaacttaataagatatagaagagagag 521
DB 301 AATATGAGAGAGAGTCTGTTGTGACGAGAGAACTTAATAGATGCTTGAAGGAAAGG 360
OY 522 aaattctgtcaatagatcattcaataaagcttcaagcttgcattgattaaagttccc 581
DB 361 AAATCTCGTCACACACTTCATTAATAAGGCTTCAGCTTGGCCCTGTGAAGTTCCC 420
OY 582 gaagacaattcttctgtatgtagcacagttataagaacaaatcatgttgaatgact 641
DB 421 GAAGCAAACTCGCTTGATGACACACAGTATTAAGCCAAATCATGTGTTGACGCTACT 480
OY 642 gttagcgtcagtaactcctgcaaggaactcacaacctatgtgtttaaagcacatat-aaa 700
DB 481 GTTGTGTCAGTACCCCTCGAGACTTATACCCCTATGTGTTTAAATGCACACATMAAA 540
OY 701 aggaagagaaacatgctcaatgagtgttctttagcaacaaagcaagagaggttaa 760
DB 541 AGGAGTGGAAACCATGTGTAAGTAGTGTGTTTAAAGCTCCAAAGCAAGAGAGGTAA 600
OY 761 actacagcacatgaattcaagagtgag-cacttttaagatcccaattgaagaatttg 819
DB 601 ACTTCAGCTCATGATGATTCAGAGGTGGAACATTTACATCTCCAACTTAAGGATCTTG 660
OY 820 gaattaagaattctctgctatataaaccacactcaagcatgatttggcaattgtg 879
DB 661 GAATTAAGATTTCTCTGCAATTAATTAATCCACCTCAGGCAATGATTCGGAATGCGG 720
OY 880 cttagagagataactgagccctgagaga-taagtaaaaaggtttgattgtgctgacag 938
DB 721 CTTCAGAGATTAACCTGATCCAGATTCATGAGACAGGCTTGTATGTGCTAGTGTG 780
OY 939 atgtc-ctgtacactcaagtgtgatalccagc--gtgtgtgagatgagcaatttgaagccag 995
DB 781 ATGTCCGCTTACACTCAGCTGTGATCATGCAAGTGTGTGATGAGACACTTGTGAGCCAG 840
OY 996 tggcctt 1001
|||||
```

DB 841 TGGCTT 846

RESULT 8
BM463514 1030 bp mRNA linear EST 05-FEB-2002
LOCUS AGENCOURT 6433730 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5585926
DEFINITION 5', mRNA sequence.

ACCESSION BM463514
VERSION BM463514.1 GI:18512556
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 1030)

AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: L1AM12352 row: P column: 23

High quality sequence start: 17

High quality sequence stop: 754.

FEATURES

source

location/Qualifiers

1..1030

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5585926"

/clone_id="NIH_MGC_92"

/tissue-type="embryonal carcinoma, cell line"

/lab-host="DH10B (phage-resistant)"

/note="Organ: testis; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dr primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH-MGC Library."

BASE COUNT 302 a 205 c 209 g 314 t

ORIGIN

Query Match 59.4%; Score 595; DB 10; Length 1030;
Best Local Similarity 100.0%; Pred. No. 1.3e-143;
Matches 595; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 407 ggcattatgcaatcaagaacacatcattatcattcattcattcattcattcattcatt 466
DB 31 GCGATTAAATGCAATCAAGCAAAACCATACCTCATATTATCTTGTATGATGTAATAT 90
OY 467 gggaaagagtttctgtgtacggaaagaaacttaataagatatagaagggagaagaaat 526
DB 91 GGGAGAGAGTTTGTGTTGTAAGGAAAGAACTTAATTAAGATTAATTAAGGAGAGCAAAAT 150
OY 527 tctctgaactgactcattacataaagctcagcttgcattgtttaaagttcccgaaag 586
DB 151 TTCTGCAATGACTTCATCATATAAAGCTTCACCTTGGCATGTTTTAAAGTTCCGAAAGC 210
OY 587 aaatcctcttgatgagacagttataagaacaaatcattgttgcattgtgc 646
DB 211 AAATCTCTTGTGATGAGACACAGTTATAAGCAAAATCATGTGTTGATGTCAGTGTG 270
OY 647 ggtcagctcctgtgagagactcacaacctatttgtttaaagacataaagaggt 706
DB 271 GGTCACTCTCTGCAAGGCTCATCAACACTATGTGTTTAAATGACATATTAAGAGAGT 330
OY 707 ggaacccattgtaatgattgttctttagcaacccaagaagaggttaactaca 766
```

|||||
Db 331 GGAACCATGTCATGATGTTGTTCTTACCAACCAAGAGAGAGGTAACACTACA 390
|||
Qy 767 gccacatgaattccagggatgacatttcacgattcccaatttagaagatttggaaattaa 826
|||
Db 391 GCCACATGTAATCCAGGGGCGACCTTTTACATCTCCAAATTTAGGATGTTGGAAATTA 450
|||
Qy 827 gaattctctgtatattaccaccccaacatgattatttggcattgtgtctcaga 886
|||
Db 451 GAATTTCTGTGATATTATTAACCACTACAGATGATTTTGGCAATTTGCTTCAGA 510
|||
Qy 887 ggaataactgtgcccgcgcagataatgaagaggttgaatgtgtcgtacgatgtctgt 946
|||
Db 511 GGATAAATGTCGCCCGCAGATATGAAAAAGGTTTGATGTGGCGATGATGCTGTG 570
|||
Qy 947 tacactcattgtgataccagggatggtgattggaagagtttggagcccatgtgctt 1001
|||
Db 571 TACACTCAATGTTGATCACCGGGGTGGTGGATGAGACGATTTGAGCCCACTGCTT 625
|||

|||||
LOCUS BG253201 857 bp mRNA linear EST 13-FEB-2001
DEFINITION 602364158f1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:4472391 5',
mRNA sequence.
ACCESSION BG253201
VERSION BG253201.1 GI:12763017
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 857)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
CONTACT: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10293 row: c column: 16
High quality sequence stop: 704.
Location/Qualifiers

|||||
FEATURES
source
1..857
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4472391"
/clone_lib="NIH_MGC_90"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: liver; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed;
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC library."
BASE COUNT 284 a 185 c 156 g 232 t
ORIGIN

Query Match 58.2%; Score 582.4; DB 10; Length 857;
Best Local Similarity 99.7%; Pred. No. 2.2e-140;
Matches 594; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 407 gcgattaagcaataaagcaacatacattatattacattcattcagatgaataat 466
|||||
Db 1 GCGATTAAAGCAATCAAGCAACCACTACTCATTTTACCTTTCTCATTAATGTAATAT 60
|||||
Qy 467 gggagaagtttctgtgtacggagaagaaactaagaatagagaaggagaagcaaat 526
|||||

|||||
Db 61 GGGAGAAGTTTGTGTGTCACGAAAGAACTTAATAAGATTTAGAAAGGAGAGCAAAAT 120
|||
Qy 527 ttctgtcaatgactcaatcaataaagcttcagcttggcattgttaaagttccgaagc 586
|||
Db 121 TTCTGTCAATGACTTCATCATTAATAAACCTTCAGCTTGGCATGTTTAAAGTTCCGAAGC 180
|||
Qy 587 aaattcttttgatgagacaggttaagaacaaatctgtgtgtgagtgatgttc 646
|||
Db 181 AAATTTCTTGTGATGAGACACAGTTATTAACAAATTCATGTTGATGATGTCAGTGTGC 240
|||
Qy 647 ggtcagtaactccgcgcagacatcacacatttggtttaatgacataaaggagt 706
|||
Db 241 GGTCACTACTCCGTCCGACGACTCATCATCACTATTTGTTTAAATGCAATATAAAGAGCT 300
|||
Qy 707 ggaacacattgctaataatgatgtgttcttcttaagcaacaaagagaggttaactaca 766
|||
Db 301 GGAACCAATGTCATTAATGATGTTGTTCTTTAGCAACCAAGAGAGAGGTAACCTACA 360
|||
Qy 767 gccacatgaattccagggatgacattttagatcccaatttagaagatttggaaattaa 826
|||
Db 361 GCCACATGTAATCCAGGGGTGGCACTTTACGATCTCCAAATTTAGAAATTTGGAAATTA 420
|||
Qy 827 gaattctctgtatattaccaccccaacatgatttggcattgtgtctcaga 886
|||
Db 421 GAATTTCTGTGATATTATTAACCACTCAGATGATTTTGGCAATTTGGCTTCAGA 480
|||
Qy 887 ggaataactgtgcccgcgcagataatgaagaggttgaatgtgtcgtacgatgtctgt 946
|||
Db 481 GGATAAATGTCGCCCGCAGATATGAAAAAGGTTTGATGTGGCGATGATGCTGTG 540
|||
Qy 947 tacactcattgtgataccagggatggtgattggaagagtttggagcccatgtgctt 1001
|||
Db 541 TACACTCAATGTTGATCACCGGGGTGGTGGATGAGACGATTTGAGCCCACTGCTT 596
|||

|||||
RESULT 10
BF939962 720 bp mRNA linear EST 22-JAN-2001
LOCUS BF939962/c
DEFINITION nacc55a04.x1 NCI-CGAP_Brn23 Homo sapiens cDNA clone IMAGE:3439014 3'
similar to SW:00D2_HUMAN P10515 DIHYDROLIPOMIDE ACETYLTRANSFERASE
COMPONENT OF PYRUVATE DEHYDROGENASE COMPLEX, MITOCHONDRIAL
PRECURSOR ; mRNA sequence.
ACCESSION BF939962
VERSION BF939962.1 GI:12357282
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 720)
NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute / National Institute of Neurological
Disorders and Stroke, Brain Tumor Genome Anatomy Project
(CGAP/BRGAP), Tumor Gene Index
Unpublished (1998)
CONTACT: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonafide, Ph.D.
DNA Sequencing by: Greg Lennon, Ph.D.
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov
Seq primer: -40up from G1bco
High quality sequence stop: 464.
Location/Qualifiers

|||||
FEATURES
source
1..720
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 284 a 185 c 156 g 232 t
ORIGIN

Db	521	CAAACCAACCATCTACTCATTTATTACCTTTCTATGCAAGTGAATATGAGCAACAGTTTCG	580
Q7	480	tttgcgcggaagaactaatatagaatattagaagcggaagaagaattctgtcaatgac	539
Db	581	-TGGTACGGAAAGAACTTAATACGATATTAGAAGGAGACAGCAAAATTTCTGTCAATGAC	639
Q7	540	ttcatcataaagaacttcagactttggcagatggtttaaagaatc	579
Db	640	tttcatcatataaaggcttcagactttggcagatggtttaaagaatc	679

RESULT	12
AA867881	
LOCUS	1125 bp
DEFINITION	AA867881 v22c10.1 Soares-thymus-2Nbdt mRNA
IMAGE:1265202	5' similar to gb:Y00978_cds1 DIHYDROLIPOAMIDE
ACTYLTRANSFERASE COMPONENT (HUMAN)''	mRNA sequence.

SESSION	AA667881	GI:2963326
SESSION	AA667881.1	
WORDS	EST.	
SOURCE	house mouse.	
ORGANISM	Mus musculus	

REFERENCE 1 (bases 1 to 1125)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,C., Soares,B., Wilson,R. and Waterston,R.
TITLE The Mashu-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project

TITLE	JOURNAL	COMMENT
The Mashu-HHMI Mouse EST Project	Unpublished (1996)	
Contact: Marra M/Mouse EST Project		
Washu-HHMI Mouse EST Project		
Washington University School of Medicine		
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108		
Tel: 314 286 1800		
Fax: 314 286 1810		
Email: mouseest@watson.wustl.edu		
This clone is available royalty-free through INL ; contact the		
IMAGE Consortium (info@image.lnl.gov) for further information.		
MC1:667754		
Seq primer: -28ml3 rev2 ET from Amersham		
High quality sequence stop:494.		

FEATURES	SOURCE	Location/Qualifiers
	1..1125	/organism="Mus musculus"
		/strain="C57BL/6J"
		/db_xref="taxon:10090"
		/clone_image:1265202"
		/clone_11b="Soares_thymus_2nbwt"
		/sex="male"
		/tissue_type="Thymus"
		/dev_stage="4 weeks"
		/lab_host="DH10B"
		/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'] TGTTACCAATCTGGAAGTCGAGACGCCCGTATTTTTTTTTTTTTTTTTT 3'; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. RNA provided by Dr. Bertrand Jordan. Library went through two rounds of normalization, and was constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT	275 a 235 c 305 g 309 t	1 others
ORIGIN		

Query Match	53.2%	Score 532.4	DB 9	Length 1125
Best Local Similarity	86.6%	Pred. NO. 2.4e-127		
Matches 609; Conservative	0	Mismatches 92;	Indels 2;	Gaps 2

[illegible]

RESULT	13				
LOCUS	BC390485				
DEFINITION	BC390485	664 bp	mRNA	linear	EST 12-MAR-2001
	60241613671 NIH_MGC_92	Homo sapiens	CDNA clone	IMAGE:452434	5',
	mRNA sequence.				
ACCESSION	BC390485				
VERSION	BC390485.1	GI:13283933			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
REFERENCE	1 (bases 1 to 664)				
AUTHORS	NIH-MGC http://mgc.ncl.nih.gov/ .				
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)				
JOURNAL	Unpublished (1999)				
COMMENT	Contact: Robert Strausberg, Ph.D.				

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be

